Access DB# 48787

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner # :_____ Date: _____

Art Unit: _____ Phone Number 30_____ Serial Number: _____

Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.
****

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

_____

Earliest Priority Filing Date: _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

. . . .

STAFF USE ONLY	Type of Search	Vendors and cost where applicable
Searcher: _Brülluki_	NA Sequence (#)_____	STN _____
Searcher Phone #: _____	AA Sequence (#)_____	Dialog _____
Searcher Location. _____	Structure (#) _____	Questel/Orbit: _____
Date Searcher Picked Up: 8/15	Bibliographic _____	Dr.Link _____
Date Completed 8/30	Litigation _____	Lexis/Nexis_____
Searcher Prep & Review Time _____	Fulltext _____	Sequence Systems _____
Clencal Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Time. _____	Other _____	Other (specify)_____

PTO-1590 (1-2000)

THIS PAGE BLANK (USPTO)

```
OM nucleic - nucleic search, using sw model

                    GenCore version 4.5
                    Copyright (c) 1993 - 2000 Compugen Ltd.

Run on:         August 29, 2001, 21:07:53 ; Search time 1467.33 Seconds
                                            (without alignments)
                                            7684.699 Million cell updates/sec

Title:          US-09-684-725-1
Perfect score:  729
Sequence:       1 atggaaaacttcagaatgc..........cactccagagtgagtatctag 729

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       1344157 seqs, 773874588 residues

Total number of hits satisfying chosen parameters:    2688314

Post-processing: Minimum Match 0%
Maximum DB seq length: 2000000000

Minimum DB seq length: 0
Maximum Match 100%
Listing first 45 summaries

Database :

    GenEmbl:*
 1: gb_ba1:*
 2: gb_ba2:*
 3: gb_ba3:*
 4: gb_in1:*
 5: gb_in2:*
 6: gb_in3:*
 7: gb_om:*
 8: gb_ov:*
 9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: gb_ba1:*
17: gb_ba2:*
18: gb_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*
44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vl:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vl1:*
59: gb_vl2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_ro1:*
95: gb_ro2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
                              %
Result                    Query
No.   Score  Match Length DB  ID        Description
----  -----  ------------ --  --------  -----------
c  1    729  100.0 216673  60  AC008571  Homo sapi
   2  717.8  98.5   1239   89  AF242874  Homo sapi
   3  717.8  98.5   1248   85  AF242874  Homo sapi
   4  717.6  98.5   1298   89  AF272363  Homo sapi
   5  716.2  98.2   1248   89  AF292402  Homo sapi
   6  471.4  64.7   1188   94  AF242875  Rattus no
   7  471.4  64.7   1314   94  AB041229  Rattus no
   8  280.8  38.5    828   93  HSOGPCR1  Homo sapi
```

```
RESULT    1
ACC0008571/c
LOCUS       AC008571    216673 bp    DNA       HTG        05-MAY-2000
DEFINITION  Homo sapiens chromosome 5 clone CTC-550M4, WORKING DRAFT SEQUENCE,
            9 unordered pieces.
ACCESSION   AC008571
VERSION     AC008571.3  GI:7708957
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 216673)
AUTHORS     DOE Joint Genome Institute.
TITLE       Sequencing of Human Chromosome 5
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 216673)
AUTHORS     DOE Joint Genome Institute.
TITLE       Direct Submission
JOURNAL     Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On May 5, 2000 this sequence version replaced gi:7211884.
            -----Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -----
            Project Information
            Center Project Name: 396672, H361
            Center clone name: CIT-HSPC_550M4
            -----
```

ALIGNMENTS

```
    9   280.8   38.5    1318     89  AF272362  Homo sapi      AF272362  Homo sapi
   10   280.8   38.5  187451     89  AC017104  Homo sapi      AC017104  Homo sapi
   11   255     35.0    1209     94  AF242873  Rattus no      AF242873  Rattus no
   12   255     35.0    1239     94  AB038649  Rattus no      AB038649  Rattus no
   13   243     33.3    1218     94  AF044602  Mus muscu      AF044602  Mus muscu
   14   159.6   21.9   75950     74  AC073449  Homo sapi      AC073449  Homo sapi
   15   134.8   18.5    2040     88  AF034632  Homo sapi      AF034632  Homo sapi
   16   134.8   18.5  163284     89  AL137000  Human DNA      AL137000  Human DNA
   17   133.2   18.3    1676     82  AF082210  Sphoeroid      AF082210  Sphoeroid
   18   127     17.4    1254     45  E11480    cDNA encodi    E11480    cDNA encodi
   19   127     17.4    4131     93  HSNEURA   H.sapiens m    HSNEURA   H.sapiens m
   20   120.4   16.5     870     97  U60181    Human growt    U60181    Human growt
   21   120.4   16.5    1101     97  HSU60179  Human growt    HSU60179  Human growt
   22   120.4   16.5  146442     74  AC069523  Homo sapi      AC069523  Homo sapi
   23   119.8   16.4  121652     60  AC008191  Drosophil      AC008191  Drosophil
   24   119.8   16.4  140838     65  AC018176  Drosophil      AC018176  Drosophil
   25   119.8   16.4  219832     60  AC007441  Drosophil      AC007441  Drosophil
   26   118.8   16.3  225374      5  AE003703  Drosophil      AE003703  Drosophil
c  27   118.8   16.3 1505606     64  AC016938  Homo sapi      AC016938  Homo sapi
   28   114     15.6     870      7  SSU60180  Sus scrofa     SSU60180  Sus scrofa
   29   114     15.6    1101      7  SSU60178  Sus scrofa     SSU60178  Sus scrofa
   30   114     15.6    1161     10  I15508    Sequence 5     I15508    Sequence 5
   31   112.4   15.4    1161     10  I15508    Sequence 2     I15508    Sequence 2
   32   112.4   15.4    1370     10  I20930    Sequence 1     I20930    Sequence 1
   33   112.4   15.4    1370     10  I20930    Sequence 1     I20930    Sequence 1
   34   112.4   15.4    1466     10  I20931    Sequence 3     I20931    Sequence 3
   35   112.4   15.4    1466     10  I28195    Sequence 17    I28195    Sequence 17
   36   112.4   15.4    1504     97  I28196    Sequence 19    I28196    Sequence 19
   37   112.4   15.4    1610     10  L12398    Homo sapien    L12398    Homo sapien
   38   112.4   15.4    1610     10  I20932    Sequence 5     I20932    Sequence 5
   39   112.4   15.1    1529     95  RNNTR2REC R.norvegicu    RNNTR2REC R.norvegicu
   40   110     15.0    1350     94  AB001982  Rattus no      AB001982  Rattus no
   41   109.2   15.0    3129     95  U94321    Rattus norv    U94321    Rattus norv
   42   106.6   14.6    3917     94  AB017027  Mus muscu      AB017027  Mus muscu
   43   101.6   13.9    2379     94  BFD2DOR   B.taurus do    BFD2DOR   B.taurus do
   44   101     13.9    1089     94  AF149717  Rattus no      AF149717  Rattus no
   45   101     13.9    2351     94  AB015645  Rattus no      AB015645  Rattus no
```

```
FEATURES             Location/Qualifiers
     source          1..216673
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone_lib="CalTech human BAC library C"
                     /clone="CTC-550M4"
BASE COUNT    66753 a    41883 c    42226 g    65000 t    811 others
ORIGIN
```

```
Summary Statistics
  Consensus quality: 197198 bases at least Q40
  Consensus quality: 207925 bases at least Q30
  Consensus quality: 210803 bases at least Q20
  Estimated insert size: 223000; pulse field gel estimation
  Estimated insert size: 215873; sum-of-contigs estimation
  Quality coverage: 5.46 in Q20 bases; pulse field gel estimation
  Quality coverage: 5.64 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1       3048: contig of 3048 bp in length
*     3049       3148: gap of unknown length
*     3149       8405: contig of 5257 bp in length
*     8406       8505: gap of unknown length
*     8506      23492: contig of 14987 bp in length
*    23493      23592: gap of unknown length
*    23593      40018: contig of 16426 bp in length
*    40019      40118: gap of unknown length
*    40119      54776: contig of 14658 bp in length
*    54777      54876: gap of unknown length
*    54877      80050: contig of 25174 bp in length
*    80051      80150: gap of unknown length
*    80151     121104: contig of 40954 bp in length
*   121105     121204: gap of unknown length
*   121205     167091: contig of 45887 bp in length
*   167092     167191: gap of unknown length
*   167192     216673: contig of 49482 bp in length.
```

```
Qy       1 atggaaaaacttcagaatgcttcctggatctaccagcagaaactagaagatccaattccag  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   84403 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG  84344

Qy      61 aaacacctgaacagcaccgaggagtatctggccttcctctgcggacctcggcgcgcagccac  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   84343 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCGCACCAC  84284

Qy     121 tcttctccccccgtgtctgtgtgtgtatgtgccaattctttgtggtggggggtcattggcaat  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   84283 TTCTTCCCTCCCCGTGTCTGTGTGTGTATGTGCCAATTTTTTGTGGTGGGGGGTCATTGGCAAT  84224

Qy     181 gtcctggtgtgctggttgattctgcagcacacaggctatgaagacgcccacacaactactac  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   84223 GTCCTGGTGTGCCCTGGTGATTCTGCAGCACACAGGCTATGAAGACGCCCACACAACTACTAC  84164

Qy     241 ctcttcaagcctgacggtctctgacctcctggtcctgctcctggaatgcccctggaggtc  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   84163 CTCTTCAAGCCTGACGGTCTCTGACCTCCTGGTCCTGCTCCTGGAATGCCCCTGGAGGTC  84104

Qy     301 tatgagaatgtggcgcaactaccccttctttcttgttcgggcccctgggtcgtgctacttcaagacg  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   84103 TATGAGAATGTGGCGCAACTACCCCTTTCTTGTTCGGGCCCCTGGGTCGTGCTACTTCAAGACG  84044

Qy     361 gccctctttgagaccgtgtgctcgctccatctcagcatcacccgcagcgtggag  420
```

Query Match   100.0%; Score 729; DB 60; Length 216673;
Best Local Similarity   100.0%; Pred. No. 4e-130;
Matches 729; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RESULT 2

```
LOCUS       AF242874    1239 bp    mRNA            PRI       13-JUL-2000
DEFINITION  Homo sapiens neuromedin U receptor 2 (NMU2R) mRNA, complete cds.
ACCESSION   AF242874
VERSION     AF242874.1  GI:9082155
KEYWORDS    .
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1239)
AUTHORS     Howard,A.D., Wang,R., Pong,S.S., Mellin,T.N., Strack,A., Guan,X.M.,
            Zeng,Z., Williams,D.L. Jr., Feighner,S.D., Nunes,C.N., Murphy,B.,
            Stair,J.N., Yu,H., Jiang,Q., Clements,M.K., Tan,C.P., McKee,K.K.,
            Hreniuk,D.L., McDonald,T.P., Lynch,K.R., Evans,J.F., Austin,C.P.,
            Caskey,C.T., Van der Ploeg,L.H. and Liu,Q.
TITLE       Identification of receptors for neuromedin U and its role in
            feeding
JOURNAL     Nature 406 (6791), 70-74 (2000)
MEDLINE     20351041
REFERENCE   2  (bases 1 to 1239)
AUTHORS     Liu,Q., McDonald,T.P., Wang,R., Jiang,Q. and Howard,A.D.
TITLE       Direct Submission
JOURNAL     Submitted (09-MAR-2000) Pharmacology, Merck Research Labs, West
            Point, PA 1946, USA
FEATURES             Location/Qualifiers
     source          1..1239
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
     gene            1..1239
                     /gene="NMU2R"
     CDS             1..1239
                     /gene="NMU2R"
                     /note="FM-4"
                     /codon_start=1
                     /product="neuromedin U receptor 2"
                     /protein_id="AAF82755.1"
                     /db_xref="GI:9082156"
                     /translation="MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAPLCGPRRSHFFLP
                     VSVVVPIFVVGNVLVCLVLICHQHAMKTPVNYLFLSLAVSDLLVLLLGMPLEVYE
                     MWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERVVAILHPFRAKLQSTRRA
                     LRILGIVWGFSVLFESLPNISHGIKFHYFPNGSLVPGSATCTVIRPMWIYNFIIQVTS
                     FLFYLLPMTVISVLVLYLMALRLKKDKSLEADEGNANIQRPCRKSVNKMLFVLVLVFAI"

BASE COUNT    267 a    392 c    265 g    315 t
ORIGIN
```

```
Query Match   98.5%;  Score 717.8;  DB 89;  Length 1239;
Best Local Similarity  99.7%;  Pred. No. 1.9e-127;
Matches  719;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

           CWAPFHIDRLFPSFVEEWSESLAAVFNLVHVVSGVFFYLSSAVNPIIYNLLSRRFQAA
           FQNVISSFPHKQWHSQOHDPQLPPAQRNIFLITECHFVELTEDIGQPFPCQSSMHNSHLPT
           ALSSEQMSRTNYQSFHFNKT"
```

```
Qy    1  atggaaaaacttcagaaatgcttcctggatctaccagcagaaaactagaagaagaatccag    60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  ATGGAAAAAACTTCAGAAATGCTTCCTGGATCTACCAGCAGAAAACTAGAAGAAGATCCAG    60

Qy   61  aaacacctgaacacgcacacggagagtatctggagctggccaatcggcggcagccac   120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  AAACAACCTTAACACACCCGAACGTTCTTGGAAGCAGCAGAAACATTCCAG   120

Qy  121  ttcttccccgtgtctgtgtgtgtatgtgccaatttttgttggtgttcattggcaat   180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  TTCTTCCCCGTGTCTGTGTGGTGTATGTGCCAAATTTTTGTTGGTGTTCATTGGCAAT   180

Qy  181  gtcctggttgtgcctggtgattctgcagacaccaggctatgaagagacgcccacaactac   240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  GTCCTGGTTGTGCCTGGTGATTCTGCAGACACCAGGCTATGAAGAGACGCCCACAACTAC   240

Qy  241  ctcttcagccctggccgtgtctctgctacctggtcctgctcctggaatgcccctggaggtc   300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  CTCTTCAGCCTGGCCGTGTCTCTGACCTCCTGGTCCTGCTCCTGGAATGCCCCTGGAGGTC   300

Qy  301  tatgagatgtggcgcaactacccctttctgttcggcgccgtgggctgctacttcaagacg   360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  TATGAGATGTGGCGCAACTACCCCTTTCTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG   360

Qy  361  gccctctttgagaccgtgtgcttcgcctccatcctcagcatcaccaccgtcagcgtggag   420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG   420

Qy  421  cgctacgtggccatcctacaccgttccgcctccgcgccaaacctgcagaacaccgggcc   480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  CGCTACGTGGCCATCCTACACCGTTCCGCCGCCAAACTGCAGACACCCGGCGCGGGGCC   480

Qy  481  ctcaggatcctcggccatcgtctgggcttccgtgtctcttctccctgccaacacagc   540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  CTCAGGATCCTCGGCAATCGTGTCGGGGCCTTCTCCGTGTCTTCTCCCTGCCAACACCAGC   540

Qy  541  atccatggcatcaagttccactacttcccccaatggggtcccaggttcggccacc   600
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCAGGTTCGGCCACC   600

Qy  601  tgtacggtgtcatcaaggcccatggtggatccaaggtcaacctcctccta   660
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  601  TGTACGGTGTCATCAAGGCCCATGGGGATCCAAGGTCAACCTCCTCCTTA   660

Qy  661  ttctacctcctccccatgactgtcatcagtgtctctactacctcaatggcactcagagtg   720
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  TTCTACCTTCCTCCCATGACTGTCATCAGTGTCTCTACTACCTCAATGGCACTCAGAGTG   720

Qy  721  agtattag   729
         ||||||||
Db  721  AGTATCTAG   83675
```

RESULT 3

```
LOCUS       AB041228    1248 bp    mRNA            PRI       21-SEP-2000
DEFINITION  Homo sapiens mRNA for G protein-coupled receptor TGR-1, complete
            cds.
ACCESSION   AB041228
VERSION     AB041228.1  GI:10257380
KEYWORDS    G protein-coupled receptor TGR-1.
SOURCE      Homo sapiens cDNA to mRNA.
```

```
Qy  721  a  721
         |
Db  721  A  721
```

ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS   Hosoya,M., Moriya,T., Kawamata,Y., Ohkubo,S., Fujii,R., Matsui,H.,
           Nishimura,O. and Fujino,M.
TITLE   Identification and Functional Characterization of a Novel Subtype
           of Neuromedin U Receptor
JOURNAL   J. Biol. Chem. 275 (38), 29528-29532 (2000)
PUBMED   10887190
REFERENCE 2 (bases 1 to 1248)
AUTHORS   Shintani,Y., Moriya,T., Ohkubo,S. and Matsui,H.
TITLE   Direct Submission
JOURNAL   Submitted (03-APR-2000) to the DDBJ/EMBL/GenBank databases. Yasushi
           Shintani, Takeda Chemical Industries, LTD, Discovery Research
           Laboratories 1; 10 Wadai, Tsukuba, Ibaraki 300-4293, Japan
           (E-mail:Shintani_Yasushi@takeda.co.jp, Tel:81-298-64-5011,
           Fax:81-298-64-5000)
FEATURES   Location/Qualifiers
     source   1..1248
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
     gene   1..1248
                    /gene="TGR-1"
     CDS   10..1248
                    /gene="TGR-1"
                    /codon_start=1
                    /product="G protein-coupled receptor TGR-1"
                    /protein_id="BAB13721.1"
                    /db_xref="GI:10257381"
                    /translation="MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLP
                    VSVVYPIFVVGVIGNVLVLCLVLLQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYE
                    MWRNYPELFGPVGCYFKTALPETVGCASILSITTVSVERYVAILHPFRAKLQSTRRA
                    IRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMIYNFIIQVTS
                    FLFYLLPMTVISVLYYLMALRLKDKSLEADEGNANIQRPCRKSVNKMLFVLVLVFAI
                    CWAPFHIDRLFFSFSVEEWSESLAAVFNLVHVVSGVFFYLSSAVNPIIYNLLSRFQAA
                    FQNVISSFHKQWHSQHDPQLPPAQRNIFLTECHFVELTEDIGPFCQSSMHNSHLPT"
BASE COUNT   269 a   393 c   269 g   317 t
ORIGIN

Best Match
Best Local Similarity 98.5%;
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

            Score 717.8; DB 85; Length 1248;
            Pred. No. 1.9e-127;

Qy   1 atggaaaactcagaatgcttcctggactctaccagcagaaaactagaagaatccaattccag 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10 ATGGAAAAACTTCAGAATGCTTCCTGGACTCTACCAGCAGAAAACTAGAAGAATCCAATTCCAG 69

Qy   61 aaacaccctgaacacgacccagggagtatctggccttcctcctgctcctcggcggcacgcac 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70 AAACAACCTGAACACGACCCAGGGAGTATCTGGCCTTCCTCCTGCTCCTCGGCGGCACGCAC 129

Qy 121 ttctcctccccgtcgtctggtgtatgtgccaatttttgtgtggtggttcattggcaat 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130 TTCTTCCTCCCCGTCGTCTGGTGTGTATGTGCCAATTTTTGTGTGGTGGTTCATTGGCAAT 189

Qy 181 gtcctggtgtgctcggtggtgtctgcagcaacaggctatgaggacgcccaccactactac 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 190 GTCCTGGTGTGCTGCGTGGTGTCTGCAGCAACAGGCTATGAAGACGCCCACCACTACTAC 249

Qy 241 ctctcagcctggccggtctctgaccctcggtcctcggcggccctgggacctggaggtc 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 250 CTCTCAGCCTGGCCGGTCTCTGACCCTCGGTCCTCGGCGGCCCTGGGAATGCCCTGGAGGTC 309

Qy 301 tatgagatgtggcgcaactaccctttctttggccgggctgtacttcaagacg 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 310 TATGAGAGTGGCGCAACTACCCTTTCTTTGTTCGGGCCCGTTGCTACTTCAAGACG 369

Qy 361 gccctctttggagaccgtgttgcttcgcctccatcctcagatcacacaccgtcagcgtgag 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ... gcccctctttggagaccgtgttgcttcgcctccatcctcagatcacacaccgtcagcgtgag 420

RESULT 4
AF272363
LOCUS   AF272363   1298 bp   mRNA   PRI   23-OCT-2000
DEFINITION Homo sapiens neuromedin U receptor 2 (NMUR2) mRNA, complete cds.
ACCESSION   AF272363
VERSION   AF272363.1 GI:10946202
KEYWORDS   .
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1298)
  AUTHORS   Raddatz,R., Wilson,A.E., Artymyshyn,R., Bonini,J.A., Borowsky,B.,
            Boteju,L.W., Zhou,S., Kouranova,E.V., Nagorny,R., Guevarra,M.S.,
            Dai,M., Lerman,G.S., Vaysse,P.J., Branchek,T.A., Gerald,C.,
            Forray,C. and Adham,N.
  TITLE   Identification and Characterization of Two Neuromedin U Receptors
            Differentially Expressed in Peripheral Tissues and the Central
            Nervous System
  JOURNAL   J. Biol. Chem. 275 (42), 32452-32459 (2000)
   PUBMED   10899166
REFERENCE   2 (bases 1 to 1298)
  AUTHORS   Bonini,J.A., Raddatz,R., Wilson,A. and Borowsky,B.
  TITLE   Direct Submission
  JOURNAL   Submitted (25-MAY-2000) Target Discovery and Assessment, Synaptic
            Pharmaceutical Corporation, 215 College Road, Paramus, NJ 07652,
            USA
FEATURES   Location/Qualifiers
     source   1..1298
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /map="5q31.1-q31.3"
     gene   1..1298
                     /gene="NMUR2"
     CDS   27..1274
                     /gene="NMUR2"
                     /note="G protein-coupled receptor"
                     /codon_start=1
                     /product="neuromedin U receptor 2"
                     /protein_id="AAG24794.1"
                     /db_xref="GI:10946203"
                     /translation="MSGMEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHF
                     VSVVYPIFVVGVIGNVLVLCLVLLQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLE
                     VEMWRNYPELFGPVGCYFKTALPETVGCASILSITTVSVERYVAILHPFRAKLQSTR
                     RALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMIYNFIIQ

Db 370 GCCCTCTTTGGAGACCGTGTTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 429
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 421 cgctacggtgccatcctaccacccgttccgccgcccaaaactgcagaacgcacccgccggggcc 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 430 CGCTACGGTGCCATCCTACCACCCGTTCCGCGCCAAACTGCAGAGCAGACCCGCCGGGGGCC 489

Qy 481 ctcaggatcctggcatcgtctggggcttcctcagtctcttcctcccctgcccaaacaccagc 540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 490 CTCAGGATCCTGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 549

Qy 541 atccatggcatcaagttccaactacttcccccaaatgggtcctctggttcggcgccacc 600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 550 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 609

Qy 601 tgtacggtcatcaagcccatgtggatctacaattcatccaaggtcacctcctccta 660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 610 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 669

Qy 661 ttctacctcctcccatgactgtcatcagtgtcctctactacctcatggccactcaagagtg 720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 670 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCCACTCAGAGTG 729

Qy 721 a 721
        |
Db 730 A 730

Left section

```
RESULT  5
        AF29402
LOCUS       AF292402     1248 bp    mRNA            PRI   11-DEC-2000
DEFINITION  Homo sapiens neuromedin U receptor-type 2 mRNA, complete cds.
ACCESSION   AF292402
VERSION     AF292402.1  GI:9944989
KEYWORDS
SOURCE      human.
```

```
        VTSFLFYLLPMTVISVLYYLMALRLKKDKSLEADEGNANIQRPCRKSVNKMLFVLVLV
        FAICWAPFHIDRLFFSFVEEWSESLAAVFNLVHVVSGVFYLSSAVNPIIYNLLSRRF
        QAAFQNVISSFHKQWHSQHDPQLPPAQRNIFLTECHFVELTEDIGPQFPCQSSMHNSH
        LPTALSSEQMSRTNYQSFHFNKT"

BASE COUNT   278 a   405 c   282 g   333 t
ORIGIN

Query Match           98.58;  Score 717.8;  DB 89;  Length 1298;
Best Local Similarity 99.7%;  Pred. No. 1.9e-127;
Matches 719;  Conservative 0;  Mismatches 2;  Indels 0;  Gaps 0;

Qy    1 atggaaaacttcagaaatgcttcctggatctaccaccagcagcagaaactagaagaagatccattccag 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   36 ATGGAAAAACTTCAGAAATGCTTCCTGGATCTACCAGCAGAAAACTAGAAGAAGATCCATTCCAG 95

Qy   61 aaacacctgaacacagaagagtatctggcctcctctgggacctcggcggcagccac 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   96 AAACACCTGAACACACGGAGGAGTATCTGGCCTTCTCTGGGACCTCGGCGGCCAGCCAC 155

Qy  121 ttcttcctcccctgtctgtcggtggtatgtgcccaattttttggtgggggtcattggcaat 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  156 TTCTTCCTCCCCTGTCTGTCGGTATGTGCCCAATTTTTGTGGTGGGGGTCATTGGCAAT 215

Qy  181 gtcctggtggtgccggtgtctgacaccccaggctatgtgcccaacaactactac 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  216 GTCCTGGTGGTGCCGGTGTCTGATTCTGCAGCACCAGGCTATGTGCCCAACAACTACTAC 275

Qy  241 ctcttcagcctggcgtctgacctcctggtctgctcctgtcctctggatgccctggaggtc 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  276 CTCTTCAGCCTGGCGTCTGACCTCCTGGTCTCCTGTGTCCTCTGGAATGCCCTGGAGGTC 335

Qy  301 tatgagatgtggcgcaactacaccctttcttcctgttcggggctgggcgtactcaagacg 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  336 TATGAGATGTGGCGCAACTACACCTTTCTTGTTGTTCGGGGCTGGGCTCTACTTCAAGACG 395

Qy  361 gcccctctttgagaccgtgtgcttcgcctcatcacacacccacgtcacgcgtgag 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  396 GCCCCTCTTTGAGACCGTGTGCTTCGCCTCATCCATCCAGCATCACCACACCGTCACGCGTGGAG 455

Qy  421 cgctacgtggccatcctacacacccgttccgcgccaaactgcagaagcaccggcgggcc 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  456 CGCTACGTGGCCATCCTACACCCGGTTCCGCGCCAAACTGCAGCACCCGGGCGGGCC 515

Qy  481 ctcaggatcctcggacatcgtctgggggctctccgtgctcttctccctgctccaacaccagc 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  516 CTCAGGATCCTCGGACATCGTCTGGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 575

Qy  541 atccatggcatcaagttccaactactttccccaatgggtcccctggttcggccacc 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  576 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTTCCCAGGGTTCGGCCACC 635

Qy  601 tgtacggtcatcaagcccatgtggatctacacaatttcatcacctcctccttccta 660
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  636 TGTACGGTCATCAAGCCCATGTGGATCTACACAATTTCATCATCCAGGTCACCTCCTTCCTA 695

Qy  661 ttctacctcctccccatgactgtcatcaagtgtcctcactacaatccaggaagtg 720
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  696 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 755

Qy  721 a 721
        |
Db  756 A 756
```

Right section

```
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1248)
AUTHORS     Shan,L., Qiao,X., Crona,J.H., Behan,J., Wang,S., Laz,T., Bayne,M.,
            Gustafson,E.L., Monsma,F.J. Jr. and Hedrick,J.A.
TITLE       Identification of a Novel Neuromedin U Receptor Subtype Expressed
            in the Central Nervous System
JOURNAL     J. Biol. Chem. 275 (50), 39482-39486 (2000)
PUBMED      11010960
REFERENCE   2  (bases 1 to 1248)
AUTHORS     Pang,L., Wang,S., Laz,T. and Hedrick,J.A.
TITLE       Direct Submission
JOURNAL     Submitted (02-AUG-2000) Human Genome Research, Schering-Plough
            Research Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033,
            USA
FEATURES             Location/Qualifiers
     source          1..1248
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /map="5"
     CDS             1..1248
                     /note="G protein coupled receptor; NmU-R2"
                     /codon_start=1
                     /product="neuromedin U receptor-type 2"
                     /protein_id="AAG03064.1"
                     /db_xref="GI:9944990"
                     /translation="MSGMEKLQNASWIYQQKLEDPFQKHLNSTEEYLAPLCGPRRSHF
                     FLPVSVVYPIFVWGVIGNVLVLVCLVILQHQAMKTPTNYLFSLAVSDLLVLLGMPLE
                     VYEMWRNYPFLFGPGVGCYFPKTALPFETVCFASILSITTVSERVVAILHPFRAKLQSTR
                     RRALRILGIVWGFSVLFSLPNTIHIGIRFHYFPNGSLVPGSARCTVIRKPMWIYNFIIO
                     VTSFLFYLLPMTVISVLYYLMALRLKKDKSLEADEGNANIQRPCRKSVNKMLFVLVLV
                     FAICWAPFHIDRLFFSFVEEWSESLAAVFNLVHVVSGVFYLSSAVNPIIYNLLSRRF
                     QAAFQNVISSFHKQWHSQHDPQLPPAQRNIFLTECHFVELTEDIGPQFPCQSSMHNSH
                     LPTALSSEQMSRTNYQSFHFNKT"

BASE COUNT   268 a   393 c   270 g   317 t
ORIGIN

Query Match           98.2%;  Score 716.2;  DB 89;  Length 1248;
Best Local Similarity 99.6%;  Pred. No. 3.9e-127;
Matches 718;  Conservative 0;  Mismatches 3;  Indels 0;  Gaps 0;

Qy    1 atggaaaacttcagaaatgcttcctggatctaccaccagcagcagaaactagaagaagatccattccag 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10 ATGGAAAAACTTCAGAAATGCTTCCTGGATCTACCAGCAGAAAACTAGAAGAAGATCCGTTCCAG 69

Qy   61 aaacacctgaacacagaacgagagtatctggcctcctctgggacctcggcggcagccac 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70 AAACACCTGAACACAGACCGAGGAGTATCTGGCCTTCTCTGGGACCTCGGCGGCCAGCCAC 129

Qy  121 ttcttcctcccccgtgtctgtggggtatgtgcccaattttttctgttggggtcaat 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  130 TTCTTCCTCCCCCGTGTCTGTGGGTGTATGTGCCCAATTTTTGTGGTGGTATGTGGCCAAT 189

Qy  181 gtcctggtgtgcctggtgatctgcagacaccaggctatgtgcccaacaactactac 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  190 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACCCCAACAACTACTAC 249

Qy  241 ctcttcagcctggcgtctgacctcctggtcctgctgggcatgcccctgggaggtc 300
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  250 CTCTTCAGCCTGGCTGTGTCTGACCTCCTGGTCCTGTTGGGCATGCCCCTGGAGGTC 309

Qy  301 tatgagatgtggcgcaactacccctttctccttctggttcctgggggcctactttcaagag 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  310 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGTCCTGGGGTGGGCTGTTACTTCAAGAG 369

Qy  361 gcccctctttgagaccgtgtgcttcgcctccatctcagcatcaccacacccgtca 420
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  370 GCCCTTCTTTGAGACCGTGTGCTTCGCCTCCATCTCAGCATCACCACACCGTCAGGGGTGGAG 429

Qy  421 cgctacgtggccatcctacaccccgttccgcgccaaactgcagaagcaccggcgggcc 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  480
```

```
RESULT 6

AF242875     1188 bp    mRNA                ROD        13-JUL-2000
LOCUS       AF242875
DEFINITION  Rattus norvegicus neuromedin U receptor 2 (NMU2R) mRNA, complete
            cds.
ACCESSION   AF242875
VERSION     AF242875.1  GI:9082157
KEYWORDS    .
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1188)
AUTHORS     Howard,A.D., Wang,R., Pong,S.S., Mellin,T.N., Strack,A., Guan,X.M.,
            Zeng,Z., Williams,D.L. Jr., Feighner,S.D., Nunes,C.N., Murphy,B.,
            Stair,J.N., Yu,H., Jiang,Q., Clements,M.K., Tan,C.P., McKee,K.K.,
            Hreniuk,D.L., McDonald,T.P., Lynch,K.R., Evans,J.F., Austin,C.P.,
            Caskey,C.T., Van der Ploeg,L.H. and Liu,Q.
TITLE       Identification of receptors for neuromedin U and its role in
            feeding
JOURNAL     Nature 406 (6791), 70-74 (2000).
MEDLINE     20351041
REFERENCE   2  (bases 1 to 1188)
AUTHORS     Liu,Q., McDonald,T.P., Wang,R., Jiang,Q. and Howard,A.D.
TITLE       Direct Submission
JOURNAL     Submitted (09-MAR-2000) Pharmacology, Merck Research Labs, West
            Point, PA 19486, USA
FEATURES             Location/Qualifiers
source               1..1188
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
gene                 1..1188
                     /gene="NMU2R"
                     /note="FM-4"
CDS                  1..1188
                     /gene="NMU2R"
                     /codon_start=1
                     /product="neuromedin U receptor 2"
                     /protein_id="AAF82756.1"
                     /db_xref="GI:9082158"
                     /translation="MGKLENASWIHDPLMKYLNSTEEYLAHLCGPKRSDLSLPVSVAY
                     ALIFLVSVMGNLLVCMVIVRHQTLKTPTNYILFSLAVSDLLVLLLGMPLEIYEMWHNY
                     PFLFGPVGCYFKTIALFVCFASILSVTTVSVERYVAIVHPFRAKLESTRRALRIIS
                     LVWSFSVVFSLPNTSIHGIKFQHFPNGSSVPGSATCSVTKPMVVNLQ0ATSFLYI
                     LPMTLISVLYLLMGLRLKRDESLEANKAVANIHRPSRKSVITMLFVLVLFAICWTPF
                     HVDRLFPSFVEEWTESLAAVNLIHVVSGVFFYLSAVNPIIYNLLSRRFRAARFNVV
                     SPTCKWCHPRHRPQGPPAQKIIFLTCHLVELETEDAGPQFPGQSSIHNTNLTTAPCAG
                     EVP"
```

```
BASE COUNT        243 a      361 c      278 g      306 t
ORIGIN

Qy     Query Match
Db     Best Local Similarity     79.8%;
       Matches    575;    Conservative    0;   Mismatches  131;    Indels   15;   Gaps    1;
```

```
Qy      1 atggaaaacttcaagaatgcttcctggatcatacagagagaaactagaagatcccaattccag      60
          ||||||||| |||||||  ||||||||||||||||||||||||||||||| ||||||| |||||
Db      1 ATGGGAAAACTTGAAAATGCTTCCTGGATCCAC--------------GATCCACTCATG      45

Qy     61 aaacaccctgaacacgaccgaggagtatctggccttctcatggacgtggcccggcagccac     120
          ||||||||||||||||||||||||||||||||||||||||||||  ||||| |||||||||
Db     46 AAGTACTTGAACAGCACAGAGGAGTACTTGGCCCACCTGTGTGGACCCAAGCGCAGTGAC     105

Qy    121 tcctcccacccgtctctgttggtgtatgtgccaatttttgtggtggggttcaattgggcaat     180
          ||||||||||| |||||||||||||||||||||| |||||| ||||||||||  |||||||
Db    106 CTATCCCTTCCGGTGTCTGTGGGCCTATGCGCTCTTCCTGGTCAGCGTGATGGGGGTAAATGGGCAAT     165

Qy    181 gtcctggtgtgcctggtgatcgtgcgcccacccaggctatgaagacgcccaccacactactac     240
          |||||||||||||||||||||||||||||||||| ||||||||||||||  |||||| |||||
Db    166 CTTCTGGTGTGCATGGTCATCGTTCGTCATCAGACATTTAAGACACCCACCAACTACTAT     225

Qy    241 ctcttcagcctggcggttctttgacctcctggtctcgctccttggaatgccctggagtc     300
          ||||||| ||||| | ||||||||||||||||||||| ||||| ||||| |||||||||| |||
Db    226 CTCTTCAGCTTGGCAGTCTCAGATCTGCTGGTGCTGCTCCTGGGCATGCCTCTGGAAATC     285

Qy    301 tatgagatgtggcgcaacactacccttcctcttgttcgggacgggtgctacttcaaagacg     360
          |||||| |||||||||||||||| ||||  |||||| |||||||| || |||| ||||| |
Db    286 TACGAGATGTGGCACAATTACCCTTTCCTGTTCGGTCCGGTGGGCTGCTACTTCAAGACA     345

Qy    361 gcccctcctttgagaccgtgctggctccctccatcctcagcgtcacacacaccgtcaggctggag     420
          ||| ||| ||||||||||||||||||| |||||||||||||  ||| ||||| || |||||
Db    346 GCCCCGTTCGAGACTGTGCTTTGCTTTGCCTCCATTCTCAGTGTCACCACGGTTAGCGTAGAG     405

Qy    421 cgctacgtggccatcctacacccgttcgcgcaaactgcagagcaccgcgcgcctggagc     480
          ||||| ||||||||| |||||| || || ||||| ||||| ||| |||||| ||| |||
Db    406 CGCTATGTGGCCATTGTCCACCCTTTCCGAGCCAAGCTGGAGAGCACCCGGAGGGCC     465

Qy    481 ctcaggagatcctcgcatccgtctgggggcttctcatcgtctctcccgtcccaacaacagc     540
          ||||| |||| ||||  ||| |||| |||||||| ||| ||||| ||| |||||||||||
Db    466 CTCAGAATCATCTCAGTCTAGTCTGGAGCTTCTCTGTGGTGTTTTCTCTCCCAATACCAGC     525

Qy    541 atccaatggcatcaagttccactacttccccaatggggtccctggtccacagttcggccacc     600
          ||||| |||| || |||| ||||| ||||||| ||| ||||| || || || |||||||||
Db    526 ATCCATGGCATCAAGTTCCAGCACTTTCCCAACGGGTCCTCCGTACCTGGCTCAGCCACC     585

Qy    601 tgtacggtcatcaagcccatgtgggatctacaatttcatcatccaggtcacccctcctccta     660
          ||||| |||||||| ||||||||||| || || |||| ||||| |||||||||| ||||
Db    586 TGCAGCGTCACCAAACCCATGTGGGTGTTAACTTGATCATCCAGGTCACCCCCTTCCTA     645

Qy    661 ttctaccctcccatgactgtgtcatcaagtgtcctctactacctcactcatggagctcagagtg     720
          |||||||||| || |||||  ||||||||||||| ||||||||||| |||||||| |||||
Db    646 TTCACACTCATCTCCGTGCTCTACTACCTCCTGATGGGGCTCCGCCTCAAGCGAGTGAC     705

Qy    721 a 721
          ||
Db    706 A 706
```

```
RESULT 7

AB041229     1314 bp    mRNA                ROD        21-SEP-2000
LOCUS       AB041229
DEFINITION  Rattus norvegicus mRNA for G protein-coupled receptor TGR-1,
            complete cds.
ACCESSION   AB041229
VERSION     AB041229.1  GI:10257382
KEYWORDS    G protein-coupled receptor TGR-1.
SOURCE      Rattus norvegicus cDNA to mRNA.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
```

REFERENCE   1 (sites)
AUTHORS   Hosoya,M., Moriya,T., Kawamata,Y., Ohkubo,S., Fujii,R., Matsui,H.,
             Nishimura,O. and Fujino,M.
TITLE   Identification and Functional Characterization of a Novel Subtype
             of Neuromedin U Receptor
JOURNAL   J. Biol. Chem. 275 (38), 29528-29532 (2000)
PUBMED   10887190
REFERENCE   2 (bases 1 to 1314)
AUTHORS   Fujii,R., Shintani,Y. and Hinuma,S.
TITLE   Direct Submission
JOURNAL   Submitted (03-APR-2000) to the DDBJ/EMBL/Genbank databases. Shuji
             Hinuma, Takeda Chemical Industries, LTD, Discovery Research
             Laboratories 1; Wadai 10, Tsukuba, Ibaraki 300-4293, Japan
             (E-mail:Hinuma_Shuji@takeda.co.jp, Tel:81-298-64-5000)
             Fax:81-298-64-5035)
FEATURES   Location/Qualifiers
     source   1..1314
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     gene   1..1314
                     /gene="TGR-1"
     CDS   1..1314
                     /gene="TGR-1"
                     /codon_start=1
                     /product="G protein-coupled receptor TGR-1"
                     /protein_id="BAA13722.1"
                     /db_xref="GI:10257383"
                     /translation="MGKLENASWIHDPLMKYLNSTEEYLAHLCGPKRSDLSLPVSVAY
                     ALIFLVGVMGNLLIVCMVIVIRHQTLKTPTNYLFSLAVSDLLVLLLGMPLEIYEMWHNY
                     PFLFGPVGCYFKTALEFTVCFASILSVTTVSVERYVAIVHPFRAKLESTRRARILS
                     LVWSFSVVFSLPNTSIHGIKFQHFPNGSSVPGSATCTVTKPMVYNLIIQATSFLFYI
                     LPMTLISVLYLLMGLRLKRDESLEANKVAVNIHRPSRKSVRTKMLFVLVLVFAICWTPF
                     HVDRLFFSFVEEWTESLAAVFNLIHVVSGVFFYLSSAVNPIIYNLLSRRFRAAFRNVV
                     SPTCKWCHPRHPQGPPAQKIIFLTECHLMELTEDAGPQFPGQSSIHNTNLTMAPCAG
                     EVP"
BASE COUNT   269 a   397 c   305 g   343 t
ORIGIN

Query Match   64.7%;   Score 471.4; DB 94; Length 1314;
Best Local Similarity   79.8%;   Pred. No. 1.9e-80;
Matches 575; Conservative   0; Mismatches 131; Indels 15; Gaps   1;

```
Qy     1 atggaaaaacttcagaatgcttcctggatctaccagcagaaactagaagatccattccag    60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   127 ATGGGAAAACTTGAAAATGCTTCCTGGATCCAC------------------GATCCACTATG   171

Qy    61 aaacacctgaacagcaccgaggagtatctgcctccatccggcgaactcggcagccac    120
         ||  ||||| ||| |||| ||||| ||| ||||  ||||| ||    ||||
Db   172 AAGTACTTGAACAGCACAGAGAGAGTACTTGGCCCACCTGTGTGGACCCAAGCCAGTGAC   231

Qy   121 ttcttcccccgtgtctggtgtatgtgcaaattttgttggttgggtcattggcaat    180
         |  |||| ||| ||||||| || ||| | ||||||||| ||||| |||||
Db   232 CTATCCCTTCCGGTGTCTGGCCTATGCCGTCAGACTTTGAAGACACCACTACTAT   291

Qy   181 gtcctggttgctgctgattgattctgcacgcaacggctatgaagacgcccacca   240
         || |||||| ||| | || ||| | ||| | | ||| |||||| || ||
Db   292 CTTCTGGTTGCTGCTATGGTTCCGACAACACAGACTTTGAAGACACCACCAACTACTAT   351

Qy   241 ctcttcaagcctgcggggtctctgacctcctggcttctctccttggaatgcccctgggagtc   300
         ||||||||| ||| |||||| || ||| ||| |||| |||||
Db   352 CTCTTCAGCTTGGCCAGTCTCAGATCTCCTGGTCCTCTGGGAAGCCTCTGGGAAATC   411

Qy   301 tatgagatgtggcgcaactacccttcttctgttcggcgccgtgggctgtgctactccaagacg   360
         || | |||| |||| |||||||| |  ||||| |||| ||||||| |||| ||||||
Db   412 TACGAGATGTGGCACAATTACCCTTTCCCTGTTCGGCGATGCTACTTCAAGACA   471

Qy   361 gccctcttttgagaacgtggtgtggttcagccccaccacacgtagcggtggag   420
         |||||| ||| |  || ||| ||| ||| ||| ||| || |||||||||||
Db   472 GCCCTCTTCGAGAACTGTGTGCTGTGCTTTGCCTCCATTCTCAGTGTCACCACCGTTAGCGTAGAG   531
```

Query Match   38.5%;   Score 280.8; DB 93; Length 828;
Best Local Similarity   65.8%;   Pred. No. 4.6e-44;
Matches 408; Conservative   0; Mismatches 212; Indels   0; Gaps   0;

```
Qy   103 ggacctcggcgcagccacactctctcctccccgtgtctgtggtgtatgtgcaatttttgtg   162
         || ||| ||| | ||| |||| ||| || ||| |||| |||||| ||||| ||||
Db    88 GGGCCCCCACCACGAGCAGAGAGTGTTCATGCCCCATCGTGTCCCACATACCTGTGTCGATCTTTCGTG   147

Qy   163 gtgggggtcattggcaatgtcctggtgtgccgtggtgattctgcagcacaggctatgaag   222
         |||||| |||||   ||  |||| ||||  | ||| ||||
```

```
Qy   421 cgctacgtggccatcctaccaccgttccggcgccaaactgcagagcacccggggc   480
         |||| |||| | |||| || || ||||||||||| ||||||||| |||| | ||| ||
Db   532 CGCTATGTGGCCATTGTCCACCCTTTCCGAGCCAAGCTGGAGAGCACGCGGCGACGGGC   591

Qy   481 ctcggaatcctcgccatcgtctggggcttctccggcgtctgctctccctgcccaaacacagc   540
         ||||||||| ||| | |||| | |||||| |||| ||| || || |||| |||
Db   592 CTCAGGATCCTCAGCGCTAGTCTGAGCGTCTGTGCTGGTCTTTGTTTGCCAATACCAGC   651

Qy   541 atccatggatcaagttccactacttcccccaaatgggtccctggttccggccac   600
         ||||||||| || ||||| || |||||| | |||| ||||| | ||| |||||
Db   652 ATCCATAGGACGATCAAGTTCCAGCACTTTCCCAACGGGTCCTCCGTACCCGGCTCAGGCACC   711

Qy   601 tgtacggtcatcaagccccatgtggatctacaaattccacagctcacctcttccta   660
         ||||||||||||||||||||||||||||| |||| ||| |||| |||| | ||
Db   712 TGCACAGTCACCAAAACCCATGTGGGTGTATAACCTTGATCATCCAAGCTACCAGCTTCCTC   771

Qy   661 ttctacctcctcccoatgactgtcatcagtgtcctcctactacccatggcaactcaagagtg   720
         ||| || || |||| | || ||| |||| |||| || || ||| |||||  || || ||
Db   772 TTCTACACTCCTCACCCTCCCAAATGACCCTCAATCACGGTCCTCTACTACCCTCACTATGTCCTC   831

Qy   721 a   721
         |
Db   832 A   832
```

```
Qy  148 GTGGGCGCTGCTGGGCAATGGGCTGCCTCGTCTGTGGTCATCCTGCGCCACAAGGCCATGCGC 207
Db  208 ACGCCTACCAACTACTACTCTTCAGCCTCTGGTCCTGCTGGTC 267
Qy  223 acgccacacaactactactcttcaagcctggcggttcctggtcctgctcctt 282
Db  268 GGGCTGCCCCTGGAGCTCTATGAGATGTGGCGCAACTACCCCATTTCCTTGCTCGGCGTTGGT 327
Qy  283 ggaatgccctcggagtctatgagatgtggcgaactaccccatttcttgttcgggcccgtg 342
Db  328 GGCTGCTATTTCCGCACGGTACTGTTTGAGATGGTCATCCTGCGCCTCAGTGTCTCAACGTC 387
Qy  343 ggctgctacttcaagacgcggcccctccttttggagaccgtgcttgccctccaatctcagcatc 402
Db  388 ACTGCCCTCGAGCGTGGAAACGCTATGTGCCGCCGTGGTGCACCACTCCACGCAGTTCATG 447
Qy  403 acaccgtcaggtgggagcgtacgttgggcactccatcaccgttccgcgcaaactgcag 462
Db  448 GTGACGCGGGGCCCCATGGTGCGCGCGAGTGCTTGGGCGGCCCTGCTCTGGGGCTTTGCCATGCTGTTGGT 507
Qy  463 agcaaccccggccgcggggccctcaggatcctccgcgcatcgtctcgtgggcttctcgtgctcttc 522
Db  508 TCCCTGCCCAACACCAGCCTGACACCTACTCCCCAATGGGTTCCCCCA 567
Qy  523 tccctgcccaacacagcacatccaagttccacctactacccccaatgggtccctg 582
Db  568 GTGCCCAGACTCAGCTGTTGCATGAGTCGGCCACCTGTCTTACACAATTGTCATCATC 627
Qy  583 gtccccaggttcggccacctgtacggtcatcaagcccatgtggattctacaatttcatcatc 642
Db  628 CAGACCAACGTCCGCCTGTCTTCTTCCTGCCGCCATCATCAAGGGCCATGCG 687
Qy  643 caggtcaacctccttcctatttctacctcccactagactgtcatcaagtgtcctcctactac 702
Db  688 CTCATTGGGCTGCGACTGCG 707
Qy  703 ctcatggcactcaagagtgag 722
```

RESULT 9

```
LOCUS       AF272362    1318 bp    mRNA            PRI    23-OCT-2000
DEFINITION  Homo sapiens neuromedin U receptor 1 (NMUR1) mRNA, complete cds.
ACCESSION   AF272362
VERSION     AF272362.1  GI:10946200
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1318)
  AUTHORS   Raddatz,R., Wilson,A.E., Artymyshyn,R., Bonini,J.A., Borowsky,B.,
            Boteju,L.W., Zhou,S., Kouranova,E.V., Nagorny,R., Guevarra,M.S.,
            Dai,M., Lerman,G.S., Vaysse,P.J., Branchek,T.A., Gerald,C.,
            Forray,C. and Adham,N.
  TITLE     Identification and Characterization of Two Neuromedin U Receptors
            Differentially Expressed in Peripheral Tissues and the Central
            Nervous System
  JOURNAL   J. Biol. Chem. 275 (42), 32452-32459 (2000)
   PUBMED   10899166
REFERENCE   2  (bases 1 to 1318)
  AUTHORS   Bonini,J.A., Raddatz,R., Wilson,A. and Borowsky,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-MAY-2000) Target Discovery and Assessment, Synaptic
            Pharmaceutical Corporation, 215 College Road, Paramus, NJ 07652,
            USA
FEATURES             Location/Qualifiers
     source          1..1318
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2q34-q37"
     gene            1..1318
                     /gene="NMUR1"
     CDS             38..1318
                     /gene="NMUR1"
                     /note="G protein-coupled receptor"
                     /codon_start=1
                     /product="neuromedin U receptor 1"
                     /protein_id="AAG24793.1"
                     /db_xref="GI:10946201"
                     /translation="MTPICLNCSVLPGDLLYPGGARNPMACNGSAARGHFPPEDLNLTD
                     EALRLKYLGLPQQTELEMPICATYLLIFVVGAVGNGLTCLVTLRHKAMRTPTNYYLFSL
                     AVSDLIVLVLGLPLELYEMWNNYPFLLGLCVLASVLNTVALSVERY
                     VAVHPLQARSMVTRAHVRRVLGAVNGLAMLCSLPNTSLHGIRQLHVPCRGPVPDSAV
                     CMLVRPRALYNMVQVQTTALLFCCLPMAIMSVYLLLIGLRLRERELLLMQEAKGRSAA
                     ARSRYTCRLQQHDRDRGRRQVTRMLFVLVVVFGICWAPFHADRVMWSVVSQWTDGLHLAF
                     QHVHVISGIFFYLGSAANPVLYSLMSSRFRETFQEALCLGACHHLRPRHSSHSLSRM
                     TTGSTLCDVGSLGSWWHPLAGNDGPEAQOETDPS"
BASE COUNT      208 a    441 c    403 g    266 t
ORIGIN
```

```
Query Match     38.5%;  Score 280.8;  DB 89;  Length 1318;
Best Local Similarity   65.8%;  Pred. No. 4.1e-44;
Matches  408;  Conservative   0;  Mismatches  212;  Indels   0;  Gaps   0;
```

```
Qy  103 ggacctcggcgcagccaacttcttctcccccgtgtctgtggttgtatgtgcaaattttgtg 162
Db  194 GGGCCCCAGCAGACAGAGCTGTTCATGCCACATCTGTGCCACATACCTGCTGATCTTCGTG 253
Qy  163 gtggggggtcattggcaaatgtcctggtgtgctggtgatctgcaacaaggctatgaag 222
Db  254 GTGGGCGGGCTGTGGGCAATGGGCTGACCTGTCTGGTCATCCTGCGCCACAAGGCCATGCGC 313
Qy  223 acgcccacacaactactactcttcaagcctggcgggtctctggaacctcggtcctgctcctt 282
Db  314 ACGCCTACCAACTACTACTCTTCAGCCTCTGGTCCTGCTGGTC 373
Qy  283 ggaatgccctgggaggtctatgagatgtggcgcaactaccccatttcttgttcgggcccgtg 342
Db  374 GGGCTGCCCCTGGAGCTCTATGAGATGTGGCGCAACTACCCCATTTCCTTGCTCGGCGTTGGT 433
Qy  343 ggctgctacttcaagacgcggcccctccttttggagaccgtgcttgccctccaatctcagcatc 402
Db  434 GGCTGCTATTTCCGCACGGTACTGTTTGAGATGGTCATCCTGCGCCTCAGTGTCTCAACGTC 493
Qy  403 acaccgtcaggtgggagcgtacgttgggcactccatcaccgttccgcgcaaactgcag 462
Db  494 ACTGCCCTCGAGCGTGGAAACGCTATGTGCCGCCGTGGTGCACCACTCCACGCAGTTCATG 553
Qy  463 agcaaccccggccgcggggccctcaggatcctccgcgcatcgtctcgtgggcttctcgtgctcttc 522
Db  554 GTGACGCGGGGCCCCATGGTGCGCGCGAGTGCTTGGGCGGCCCTGCTCTGGGGCTTTGCCATGCTGTTGGT 613
Qy  523 tccctgcccaacacagcacatccaagttccacctactacccccaatgggtccctg 582
Db  614 TCCCTGCCCAACACCAGCCTGACACCTACTCCCCAATGGGTTCCCCCA 673
Qy  583 gtccccaggttcggccacctgtacggtcatcaagcccatgtggattctacaatttcatcatc 642
Db  674 GTGCCCAGACTCAGCTGTTGCATGAGTCGGCCACCTGTCTTACACAATTGTCATCATC 733
Qy  643 caggtcaacctccttcctatttctacctcccactagactgtcatcaagtgtcctcctactac 702
Db  734 CAGACCAACGTCCGCCTGTCTTCTTCCTGCCGCCATCATCAAGGGCCATGCG 793
Qy  703 ctcatggcactcaagagtgag 722
Db  794 CTCATTGGGCTGCGACTGCG 813
```

RESULT 10

```
LOCUS       AC017104   187451 bp    DNA             HTG    16-MAR-2001
DEFINITION  Homo sapiens chromosome 2 clone RP11-562I5, WORKING DRAFT SEQUENCE,
```

```
ACCESSION   AC017104
VERSION     AC017104.6  GI:13357551
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 187451)
AUTHORS     Waterston,R.H.
TITLE       The sequence of Homo sapiens clone
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 187451)
AUTHORS     Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (09-DEC-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Mar 16, 2001 this sequence version replaced gi:9211494.

            -------------- Genome Center --------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc/index.shtml
            ------- Project Information -------
            Center project name: H_NH0562105
            -------- Summary Statistics --------
            Sequencing vector: M13; 37%
            Sequencing vector: plasmid; 63%
            Chemistry: Dye-primer ET; 37% of reads
            Chemistry: Dye-terminator Big Dye; 63% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 184599 bases at least Q40
            Consensus quality: 185465 bases at least Q30
            Consensus quality: 185902 bases at least Q20
            Insert size: 186551; sum-of-contigs
            Insert size: 194000; agarose-fp
            Quality coverage: 7.91 in Q20 bases; sum-of-contigs
            Quality coverage: 7.35 in Q20 bases; sum-of-contigs

            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 10 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.

            *  1     1504: contig of 1504 bp in length
            *  1505  1604: gap of unknown length
            *  1605  8409: contig of 6805 bp in length
            *  8410  8509: gap of unknown length
            *  8510  16888: contig of 8379 bp in length
            *  16889 16988: gap of unknown length
            *  16989 22115: contig of 5127 bp in length
            *  22116 22215: gap of unknown length
            *  22216 31817: contig of 9602 bp in length
            *  31818 31917: gap of unknown length
            *  31918 44446: contig of 12529 bp in length
            *  44447 44546: gap of unknown length
            *  44547 69386: contig of 24840 bp in length
            *  69387 69486: gap of unknown length
            *  69487 97692: contig of 28206 bp in length
            *  97693 97792: gap of unknown length
            *  97793 126980: contig of 29188 bp in length
            *  126981 127080: gap of unknown length
            *  127081 187451: contig of 60371 bp in length

FEATURES             Location/Qualifiers
source               1..187451
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /clone="RP11-56215"

10 unordered pieces.
misc_feature    1..1504
                /note="assembly_name:Contig10"
misc_feature    1605..8409
                /note="assembly_name:Contig11"
misc_feature    8510..16888
                /note="assembly_name:Contig12"
misc_feature    16989..22115
                /note="assembly_name:Contig13"
misc_feature    22216..31817
                /note="assembly_name:Contig14"
misc_feature    31918..44446
                /note="assembly_name:Contig15"
misc_feature    44547..69386
                /note="assembly_name:Contig16"
misc_feature    69487..97692
                /note="assembly_name:Contig17"
misc_feature    97793..126980
                /note="assembly_name:Contig18"
misc_feature    127081..187451
                /note="assembly_name:Contig19"

BASE COUNT    47850 a  45294 c  45833 g  47569 t   905 others
ORIGIN

Query Match     38.5%;  Score 280.8;  DB 64;  Length 187451;
Best Local Similarity  65.8%;  Pred. No. 1.2e-44;
Matches  408;  Conservative  0;  Mismatches 212;  Indels  0;  Gaps  0;

Qy    103 ggacctcggcgcagccacttcttcctcccgtgtctgtggtgtatgtgccaatttttgtg 162
          |||||  | ||| | || || |  || | |||| || | | | |  | ||||  |||| |||
Db 146088 GGGCCCCAGCAGCAGAGCCTGTTCATGCCCATCTTGTGCCACATACCTGCTGAATCTTCGTG 146147

Qy    163 gtggggtcattggcaatgtcctggttgtcctggttgattctgcagcaccaggctatgaag 222
          ||||| | | || || ||| | | | | || ||  | |  |  || ||  || || | ||||
Db 146148 GTGGGCGGGCTGTGGGCAATGGGCTGACCGTGTCTGGTCATCCGTCGCCCACAAGGCCATGGCC 146207

Qy    223 acgcccaacaactactactctctcagcctggccgggtcttctggacctcctggtctccctt 282
          | |  || ||  | | |  | |  ||| |  || ||  | |  || | | || | | ||||
Db 146208 ACGGCTACCAACTACTACCTCTTCAGCCCTGGCCTGTGTCCGACCTGCTCGGTGCTCGGTG 146267

Qy    283 ggaatgccctggaggttctatgagagtgtggcgcaactaccccttctcttgttcgggcccgtg 342
          || | ||| ||| | |  | |||  | || || ||| |  |||  || ||  | |  |  |||
Db 146268 GGCCTGCCCCTGAGCTCTATGAGATGTGGCCACAACTACCCCTTCTCCTGTCTGGGCGTTGGT 146327

Qy    343 ggctgctacttcaagacgcgcccctctttgagacgtgtgcttcgccctccatcctcagcatc 402
          || || ||  | ||| | | |  |||  | ||| | | || | || | | || ||| || |
Db 146328 GGCTGCTATTTCCCCACGTGTTTTGAGATGGTTCCGGCCTCAGTGCTCACTGTGCTCAAACGTC 146387

Qy    403 acaacgtcaagcgtggagcggctacgtgtgccatcctacacccgttccggcgccaaactgcag 462
          | | || ||| || | || |  |  | || | |  | |  || |  | ||| ||  | |||||
Db 146388 ACTGCCCCTGAGCGTGGAAGCTATGTGGGCCGTGGTGCACCCACTTCCAGGCCAGGTGCCATG 146447

Qy    463 agcacccggcgccggggccctcaggatccteggcaatcgtctggggcttctccgtgctcttc 522
          | || |  ||  | ||| || | ||| ||| |||  |  || || |||||  | || | | |
Db 146448 GTGACGCGGGCCCCATGTCCGCCGAGTGCTTGTTGGGGGTCTTTGCCATGCTGTGC 146507

Qy    523 tccctgcccaacacagcatccatggcatcaagttccactacttcccccaaatgggtccctg 582
          |  || || ||  | | | || || |  || ||| | |  | |  || | |  ||| ||||
Db 146508 TCCCTGCCCAAACACAGCCATCCGGCAGTCCACGTGGGCCGCTCCACGTGCCAGGTGCCA 146567

Qy    583 gtcccaggttcggcaacctgtagggtcatcaagcccacatgtgggatcacaaattcatcatc 642
          || || |||| ||  || | ||  ||| ||||| || | |||| ||  |||||| | |||||
Db 146568 GTGGCCGAGACTCAGGTGTTTGCAATGCCTGGTCCGCCGCCACCGGCCCCTTCTACGAGTGG 146627

Qy    643 caggtcacctcttccttctacctcctcccccatgactgtcatcagtgtcctctactac 702
          || ||| | | ||| | | | | || | |  | | | | |||  | || | |||  | ||
Db 146628 CAGACCACCGCGCTGCTCTTCTTCTTCCTGCCCATGCCATCATGAGAGCCTGCGCTACCTG 146687

Qy    703 ctcatggcactcagaatgag 722
          || | || | |  ||  |
Db 146688 CTCATTGGGCTGCCGACTGCG 146707
```

RESULT 11
```
LOCUS       AF242873      1209 bp    mRNA      ROD    13-JUL-2000
DEFINITION  Rattus norvegicus neuromedin U receptor 1 (NMU1R) mRNA, complete
            cds.
ACCESSION   AF242873
VERSION     AF242873.1  GI:9082153
KEYWORDS    .
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1209)
AUTHORS     Howard,A.D., Wang,R., Pong,S.S., Mellin,T.N., Strack,A., Guan,X.M.,
            Zeng,Z., Williams,D.L. Jr., Feighner,S.D., Nunes,C.N., Murphy,B.,
            Stair,J.N., Yu,H., Jiang,Q., Clements,M.K., Tan,C.P., McKee,K.K.,
            Hreniuk,D.L., McDonald,T.P., Lynch,K.R., Evans,J.F., Austin,C.P.,
            Caskey,C.T., Van der Ploeg,L.H. and Liu,Q.
TITLE       Identification of receptors for neuromedin U and its role in
            feeding
JOURNAL     Nature 406 (6791), 70-74 (2000)
MEDLINE     20351041
REFERENCE   2  (bases 1 to 1209)
AUTHORS     Liu,Q., McDonald,T.P., Wang,R., Jiang,Q. and Howard,A.D.
TITLE       Direct Submission
JOURNAL     Submitted (09-MAR-2000) Pharmacology, Merck Research Labs, West
            Point, PA 19486, USA
FEATURES             Location/Qualifiers
     source          1..1209
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
     gene            1..1209
                     /gene="NMU1R"
     CDS             1..1209
                     /gene="NMU1R"
                     /note="FM-3"
                     /codon_start=1
                     /product="neuromedin U receptor 1"
                     /protein_id="AAF82754.1"
                     /db_xref="GI:9082154"
                     /translation="MSCNDSEFKEHFDLEDLNLTHEDLRLKYLGPQQVKQFLPICVTY
                     LIIFVVGTIGNGLTCTVILRQKAMHTPTNFYLFSLAVSDLLVLLVGLPLELYEMQHNY
                     PFQLGAGGCYFRILLLETVCLASVLNVTALSVERYVAVVHPLQAKSVMTRHVRRMLG
                     AIWPAILFSLPNTSLHGISPLIYPCRGPVPDSVTCTLVRPQFFYKLVIQTTILLFFC
                     LPMTVISVLLILGLRLRRERIILLQEEVKGRISAAAAROASHRSIOLRDRERRQVTKML
                     IALVIVFGTVCWPPHADRLMWSWVSHWTDGLRLAFQSVHLASGVFLIGSAANPVLYN
                     LMSTRFRESFRETLGIGTRCCHRHQPRHDSHSHLRLTTVSTLCDRNSRDVPLAENRDP
                     GCEGEDPPE"
BASE COUNT      239 a    347 c    323 g    300 t
ORIGIN
```

Query Match 35.0%; Score 255; DB 94; Length 1209;
Best Local Similarity 62.8%; Pred. No. 3.5e-39;
Matches 396; Conservative 0; Mismatches 235; Indels 0; Gaps 0;

```
Qy   99 ctgcggacctcggggacgcacttcttcctcccgtgtctgttggtgtatgtgcccaatttt 158
         ||  ||| |||||||| |||||| || | ||| ||  ||| ||||| ||  ||||| |||
Db   84 CTTGGGGCCCACAGCGGTAAAACAATTTTTGCCCATCTGTGTCACGTACCTGTTGATCTT 143

Qy  159 tgtggtgggtcattggcaaatgtcctggtgtgcctggtgattctgagacaccaggctat 218
         ||  || |  || ||||| |||  || ||||| |||  | | | ||| || || |||
Db  144 CGTAGTGGGCCACTCTGGCCAACCGGTTGACCTCCACGGTCATCCTGCGCCAGAAGCAAT 203

Qy  219 gaagagcgcccaccaaactacatactctccagcctggcggtctcctgacctctggtcctgct 278
         |||||  || ||| || |||| ||||||||||| || |||||| || ||  ||| |||
Db  204 GCACACGCCCACCAAACTTCTACCTCTTCAGTCCTGCGGTGTCCGATTGGTGGTCGTCCT 263

Qy  279 ccttggaatgccctgaggtctcataagatgtggcgcaactaccacccttcttcgttcgggcc 338
         |  || ||  |||||| || ||| | | ||  | |||| |||| ||||| | | ||| ||
Db  264 GGTGGGCGTTGCCCTGAACTTTATGAGATCCAGCACAATTACCCATTCCAGCTGGGGTGC 323
```

RESULT 12
```
LOCUS       AB038649      1239 bp    mRNA      ROD    11-JUL-2000
DEFINITION  Rattus norvegicus mRNA for G protein-coupled receptor FM-3,
            complete cds.
ACCESSION   AB038649
VERSION     AB038649.1  GI:9049392
KEYWORDS    .
SOURCE      Rattus norvegicus
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (sites)
AUTHORS     Fujii,R., Hosoya,M., Fukusumi,S., Kawamata,Y., Habata,Y.,
            Hinuma,S., Onda,H., Nishimura,O. and Fujino,M.
TITLE       Identification of neuromedin U as the cognate ligand of the orphan
            G protein-coupled receptor FM-3
JOURNAL     J. Biol. Chem. 275 (28), 21068-21074 (2000)
MEDLINE     20347213
REFERENCE   2  (bases 1 to 1239)
AUTHORS     Fukusumi,S., Fujii,R. and Hinuma,S.
TITLE       Direct Submission
JOURNAL     Submitted (22-FEB-2000) to the DDBJ/EMBL/GenBank databases. Shuji
            Hinuma, Takeda Chemical Industries, LTD, Discovery Research
            Laboratories 1; Wadai 10, Tsukuba, Ibaraki 300-4293, Japan
            (E-mail:Hinuma_Shuji@takeda.co.jp, Tel:81-298-64-5035,
            Fax:81-298-64-5000)
FEATURES             Location/Qualifiers
     source          1..1239
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     gene            1..1239
                     /gene="FM-3"
     CDS             1..1239
                     /gene="FM-3"
                     /codon_start=1
                     /product="G protein-coupled receptor FM-3"
                     /protein_id="BAA99387.1"
                     /db_xref="GI:9049393"
                     /translation="MLSPNASTGLLSCNDSEFKEHFDLEDLNLTHEDLRLKYLGPQQV
                     KQFLPICVTYLLIFVVGTIGNGLTCTVILRQKAMHTPTNFYLFSLAVSDLLVLLVGLP
                     LELYEMQHNYPFQLGAGGGCYFRILLLETVCLASVLNVTALSVERYVAVVHPLQAKSVM
```

```
Qy  339 cgtgggtgctactcaagacggacgggcccctcttttgagacgtgtgcttcgcctccaatccag 398
         ||  |||   |||| |||||   |||||||| ||| ||||| ||| ||| || ||  |||
Db  324 AGGTGGCGTGTTACTTCCGGATACTTGCTTTTGGAGACTGTCGTCTGCCTGGCTCAGTCGTCAA 383

Qy  399 catcaaccaccgtcaagctgggacgctacggtggccatcctacaaccgttccgcgcaaact 458
         ||| ||| || || ||| || ||||| |||||||||||  | ||||| || || ||
Db  384 TGTCACACCCCTAAGTGTGGACGGTTATGTGGCCGTGGTGCACCCACTCCAAGCCAAGTC 443

Qy  459 gcagagcaccccggcgccggcccctcaggatcctcggcgatcgtctggggcttctcctgct 518
         ||  || || ||  || ||||  ||||| | |  ||| || ||||||| | |||| |||
Db  444 TGTGATGACACGGCACGTCCCCGGCGCCGGCCATCTGGGAGCCATCTGGCCGGTCTTCGCTTATTCT 503

Qy  519 cttctccctgcccaaacaccagcatccatggccactcaagttccaactacttcccaatggtc 578
         ||||| |||||||||  |||||| | ||||||| |||| |||||| ||||| || |||  ||
Db  504 CTTCTCTCTGCCCAACACCAGCTTACATGGCCTCAGTCCACTCATATGTAACCTGCGGG 563

Qy  579 cctggtcccaaggttcggcccacctgtacggtcatcaaagcccatggatctacaatttcat 638
         ||| | |||  ||||  || | |   || ||||||||  ||||| || ||||||  | ||
Db  564 GCCGGGTGCCCCATTCAGTTACGTGGTGCGTGGTGGGTGCCCCAGTTCTTCTACAAGTTGGT 623

Qy  639 catccaggtcaccctccttctattctcctcccgatgactgtcatcagtgtcctcta 698
         ||  | || |||| || | || || ||||||| ||||| |||||||||||  | |
Db  624 AATACAGACGACGACCATACTGCTCTTCGTTGTTCTCGGCCATGTGACCATCAGTGTGCCTGTA 683

Qy  699 ctacctccatggcactccaagagtgagtatctcag 729
         |  ||||| ||||||| ||||||||||| |||||
Db  684 CCTGCTCATTGGGCTGGAGGCTGCGGAGGGAG 714
```

```
TITLE        Direct Submission
AUTHORS      Hrenuk,D.L., Smith,R.G. and Howard,A.D.
             McKee,K.K., Tan,C.T., Liu,J., Palyha,O.C., Feighner,S.D.,
REFERENCE    1  (bases 1 to 1218)
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ORGANISM     Mus musculus
SOURCE       house mouse.
KEYWORDS     .
VERSION      AF044602.1  GI:2865471
ACCESSION    AF044602
DEFINITION   Mus musculus orphan G protein-coupled receptor mRNA, complete cds.
LOCUS        AF044602                1218 bp    mRNA            ROD       11-FEB-1998
```

```
FEATURES             Location/Qualifiers
source          1..1218
                /organism="Mus musculus"
                /db_xref="taxon:10090"
CDS             1..1218
                /codon_start=1
                /product="orphan G protein-coupled receptor"
                /protein_id="AAC02681.1"
                /db_xref="GI:2865472"
                /note="similar to receptors for growth hormone
                secretagogues and neurotensin"
                /translation="MVCNISEFKWPYQPEDLNITDEALRLKYLGPQQMKQFVPICVIY
                LLIFVVGTLGNGLTCTVLLRNKMRTPTPNFYLFSLAVSDMLLVLLVGLPLELYEMQQNY
                PFOLGASACYFRILLLEETVCLASVLNVTALSVERYVAVVRPLQAKSYMTRAHVRRMVG
                AIWVLATLFSLPNTSLHGLSQLIVPVPCRGVPVPDSAICSLVGPMDFYKLVVLTFALLFPC
                LPMVTISYLLILGLRLRERRMLLQOEVKGRKTAATOETSHRRIQLQDRGRQVTKML
                FALVVVGICWAPFHADRIMWSLVYGHSTEGLHLAYQCVHIASGIFYLGSAANPVLY
                SLMSTRFREETFLQALGLGTOCCHRROPYHGSHNHIRLTTGSTLCDVGHRNSRDEPLAV
                NEDPGCQOETDPS"
ORIGIN
```

```
BASE COUNT   238 a    363 c    334 g    283 t

Query Match  Best Local Similarity  33.3%;  Score 243;  DB 94;  Length 1218;
Matches 387; Conservative  61.7%;  Pred. No. 6.8e-37;  Mismatches 240;  Indels  0;  Gaps  0;
```

```
Qy    99 ctgcgacctcgggcgcgcagccacctcttcctccccgtgtcttgtggtgtatgtgccaatttt 158
         | |||   ||||||| || ||| |  ||   ||||||| || | |||| ||  ||  ||
Db   114 CTTGGGGCCACAGCAGGTAAAACAATTTTTCCCCATCTGTGTCACGTACCGTTGATCTT 173

Qy   159 tgtggtgggggtcattggcaatgtccctggttggtgtgccctggttgattctgaagcaccagcctat 218
         || ||    |||| ||||| |||  | |||| ||| ||    |||| || | || || |
Db   174 CGTAGTGGGCACTCTGGGCAACGGGTTGACCTGCACGTCATCATCTCGGCCAGAAGGCAAT 233

Qy   219 gaagacgcccaccaactactaaccctcttcagcctgtggggtctctgacctcggtctgct 278
         | || || ||| ||||||    |||| |||||   || |||||||||||  || ||||| ||
Db   234 GCACACCCCACCAAACTTCTACCTCTTCAGTCTGCCGGTGTCCGGATTTGCTGGTGTGCTCCT 293

Qy   279 ccttggaatgcccctggaggtctatatgagaatgtggcggcaactacccttttcttgttcgggcc 338
         | |||| || |   ||| ||||   |  || || |  |||  | ||| || ||  |||  ||||
Db   294 GGTGGGCTTGCCCCTGGAACTTTATGAGATGCAGCACAATTACCCATTCCAGCTGGGTGC 353

Qy   339 cgtggggtgctacttcaagagacgctgtctttgctcgcctccatcctcag 398
         ||| |   ||||| ||| || |  |||  |  |||| ||||| ||| || |
Db   354 AGGTGGCGTTGACTTCCGGATAACTGCTTTGGAGACTGTCGCCTGGCTTCAGTGCTCAA 413

Qy   399 catcacaacgtgcaagacgctacggtggagcgtcacctacaacaccgttccgcgccaaact 458
         |||||| |||||  |||  ||||  ||| || ||| |  ||  || || ||||| ||| |
Db   414 TGTCACAGCCCTTAAGTGTGTGGGAACCGTTATGTGTGCACCCAGTTCAAGGACCATATGGCC 473

Qy   459 gcagagacgacccggcgcgcggccctcaggaatcctcggcaatcgtctggggcttctccgtgct 518
         |  || ||| ||| || || ||||| || || ||| || || | |||| ||  || ||| || |
Db   474 TGTGATGACACGGACCCCATGTCGCCGCCATCTGGGTGTGCCGGTGTCCGATAAGCTGGTTGCT 533

Qy   519 cttctccctgcccaaacaccagatcaagttccactcactacctcccccaatgggtc 578
         |  |||||| ||| ||  |||   || |||| || ||| ||   | |||||   |
Db   534 CTTCTCTGCTGGCCCAACACCAGCTTACAGGGCCTCACTCTATGTACCTGGGGG 593

Qy   579 cctggtcccaggttcgggccacctgtacggtcatcaagcaccccctcttgttgcttcgcctcatcctcagcatc 638
         || ||| ||||| |   |  || ||  ||| ||||| |||||    || ||||  |  |||| || ||
Db   594 GCCGGGTGCCCGATTCAGTTCAGTTACGCTGGTGCGTCCCAGTTCTTCTGGTCCTTCAATGTC 653

Qy   639 catcccaggtcacctctcttcattctcaccccatgactgtcatcagtgtcctcta 698
         ||   ||  |   |  ||| | ||  ||| |||| ||  ||| |  |||| |||
Db   654 AATACAGACGACGACCATACTGCTCTTCTTGTCGTCCCATGGTCACCATCACTGTGCTGTA 713

Qy   699 ctacctcatggaactcagagtgaagtatactag 729
         |  |||  ||||||||  ||||  ||| | |||
Db   714 CCTGCTCATTGGGCTGAGGCTGCGGGAGGAGGAG 744
```

(lower / second column)

```
BASE COUNT   245 a    359 c    329 g    306 t
ORIGIN

Query Match  Best Local Similarity  35.0%;  Score 255;  DB 94;  Length 1239;
Matches 396; Conservative  62.8%;  Pred. No. 3.5e-39;  Mismatches 235;  Indels  0;  Gaps  0;
```

```
Qy   103 ggacctcgggcgcgcagccacctcttcctccccgtgtcttgtggtgtatgtgccaatttttgtg 162
         || |||||| ||||  | |||||||||||| |||||| ||  |||| ||  || |||
Db    88 GGGCCACAGCCAGATGAAAACAGTTTGTCCCCATCTGTGTCACGTACCGTTGATCTTCGTG 147

Qy   163 gtgggggtcattggcaatgtccctggttgtgtcttgctcttgattctgaagcaccaggctatgaag 222
         |||| ||| |||||   |||  | || ||||| |   ||||| |||||  |||| | |
Db   148 GTGGGCACTCTGGGCAACGGGTTGACCTGCACGTCATCATCTCGGCCAACAAGACTATGGCC 207

Qy   223 acggccacaaactactaccctcttcaagcctggcgggtctctgacctcggtctgct 282
         ||||| |||||||| || |||||||| ||| | |||  |||||||| || |||| |
Db   208 ACGCCCACCAAACTTCTACCTCTTCAGCCTGCTCCGCGTGTCCGGATATGCTGGGTGTCCTCCTGTGG 267

Qy   283 ggaatgccctggaggtctatgagatggcggcaactacccttttcttgttcgggccccgtg 342
         ||||| ||||| |  ||| |||||| |  | ||||| |   |||  |||| || |||
Db   268 GGCTTGCCCTCTGGAGCTTTATGAGATGCAGCAAAATTACCCGTTCCAGCTGGGTGCGAGT 327

Qy   343 ggtgctacttcaagacgcgccctctttgagacggtgtgctttgcctccatcctcagcatc 402
         ||| ||||| |||  |  |||||| || || || |||| || |||| || || || |||
Db   328 GCCTGCTACTTCCGAATTACTGCTTTAGAGACCGTTCGCCTAGCTTCAGTGCTCAAATGTC 387

Qy   403 accaccgtcaagctggagcgtacggtggccatcctacaaccggttccggccaaactcag 462
         ||||| || | |||  ||| || |  || ||| |  ||  || ||| || ||  |||
Db   388 ACAGCCCTTAAGTGTGTGGGAACCGTTATGTGGCCATCCTACAACCGTCCGGGCAAACTGCAG 447

Qy   463 agcaccgcgccgggccctcaggatcctcgggcatcgtctggggcttctctgctctc 522
         || ||| | || || ||| || || ||| ||  | | ||||  ||||| | || |
Db   448 ATGACACGGACCCCATGTCGCCGCCATCTGGGTGTGCCGGTGTCCGATAAGCTGGTTGCTACTCTTC 507

Qy   523 tcctgcccaaacaccagatcaagttccactacttcccccaatgggttcctg 582
         |||||| ||  |||||  || |||| || |  ||   | ||||  | || |
Db   508 TCTCTGCCAACACCAGCCTTGCATGGCCTCAACTAACTGTGCCCTGGGGGCCG 567

Qy   583 gtcccaggttcgggccacctgtacggtcatcaagcacccatggtggatctacaaattcatcatc 642
         ||| |||| |   |  || || | |||| || ||||| |  ||| || || | ||| ||||
Db   568 GTCGCCCGGACTCAGTTCAGTTATAGTGGTGCGTCCCAGTTCTTCTGGTCCTTCAATGTGGTA 627

Qy   643 caggtcacctctccttattctaactcctcccccatgactgtcatcagtgtcctcctactac 702
         ||| | |    ||| | |||| ||| ||  ||||| ||  ||| |  |||| | || ||
Db   628 CTGACTACGGCACTCGCTCTTCTTGTCGTCCCATGGTCACCATCAGTGTCACCATATATCG 687

Qy   703 ctcaatggactcagagtgaagtatactag 729
         ||  |  |  ||||||||  ||| | |||
Db   688 CTCATTGGGCTGCGGCTGCGGGAGGAGGAG 714
```

```
TRTHVRRMLGAIWFAILFSLPNTSLHGLSPLYVPCRGPVPDSVTCTLVRPQFFYKLV
IQTTILLFPCLPMVVTISVLLLIGLRLRRERMLLQEEVKGRISAAARQASHRSIQLRD
RERQVTKMLIALVIVFGTCWVPFHADRLMWSVSHWTDGLRLAFQSVHLASGVFLYL
GSAANPELYNLMSTRFRESERETLGLGTRCCHRHQPHRDSHSHLRLTTVSTLCDRNSR
DVPLAENRDPGCEQETDPPE"
```

```
RESULT  14
LOCUS       AC073449/c
AC073449/c
DEFINITION  Homo sapiens clone RP11-154A7, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC073449
VERSION     AC073449.1  GI:8571740
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 75950)
AUTHORS     Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE       Homo sapiens, clone RP11-154A7
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 75950)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastian,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
            Direct Submission
JOURNAL     Submitted (18-JUN-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     All repeats were identified using RepeatMasker;
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            --------------------- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            --------------------- Project Information
            Center project name: L10390
            Center clone name: 154_A_7
            -----------------------------------------------------
            * NOTE: This record contains 92 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *
            *        1      733: contig of 733 bp in length
            *      734      833: gap of 100 bp
            *      834     1540: contig of 707 bp in length
            *     1541     1640: gap of 100 bp
            *     1641     2334: contig of 694 bp in length
            *     2335     2434: gap of 100 bp
            *     2435     3156: contig of 722 bp in length
            *     3157     3256: gap of 100 bp
            *     3257     3981: contig of 725 bp in length
```

```
*     3982     4081: gap of 100 bp
*     4082     4828: contig of 747 bp in length
*     4829     4928: gap of 100 bp
*     4929     5658: contig of 730 bp in length
*     5659     5758: gap of 100 bp
*     5759     6493: contig of 735 bp in length
*     6494     6593: gap of 100 bp
*     6594     7329: contig of 736 bp in length
*     7330     7429: gap of 100 bp
*     7430     8169: contig of 740 bp in length
*     8170     8269: gap of 100 bp
*     8270     8974: contig of 705 bp in length
*     8975     9074: gap of 100 bp
*     9075     9806: contig of 732 bp in length
*     9807     9906: gap of 100 bp
*     9907    10636: contig of 730 bp in length
*    10737    11426: gap of 100 bp
*    11427    11526: contig of 690 bp in length
*    11527    12261: gap of 100 bp
*    12262    12361: contig of 735 bp in length
*    12362    13097: gap of 100 bp
*    13098    13197: contig of 736 bp in length
*    13198    13938: gap of 100 bp
*    13939    14038: contig of 741 bp in length
*    14039    14779: gap of 100 bp
*    14780    14879: contig of 741 bp in length
*    14880    15624: gap of 100 bp
*    15625    15724: contig of 745 bp in length
*    15725    16471: gap of 100 bp
*    16572    17304: contig of 747 bp in length
*    16572    17304: gap of 100 bp
*    17305    18113: contig of 733 bp in length
*    17405    18213: gap of 100 bp
*    18114    18940: contig of 709 bp in length
*    18214    18940: gap of 100 bp
*    19041    19753: contig of 727 bp in length
*    19754    19853: gap of 100 bp
*    19854    20551: contig of 713 bp in length
*    20552    20651: gap of 100 bp
*    20652    21378: contig of 698 bp in length
*    21379    21478: gap of 100 bp
*    21479    22207: contig of 727 bp in length
*    22208    22307: gap of 100 bp
*    22308    23029: contig of 729 bp in length
*    23030    23129: gap of 100 bp
*    23130    23872: contig of 722 bp in length
*    23873    23972: gap of 100 bp
*    23973    24699: contig of 743 bp in length
*    24700    24799: gap of 100 bp
*    24800    25534: contig of 727 bp in length
*    25535    25634: gap of 100 bp
*    25635    26371: contig of 735 bp in length
*    26372    26471: gap of 100 bp
*    26472    27211: contig of 737 bp in length
*    27212    27311: gap of 100 bp
*    27312    28018: contig of 740 bp in length
*    28019    28118: gap of 100 bp
*    28119    28842: contig of 707 bp in length
*    28843    28942: gap of 100 bp
*    28943    29675: contig of 724 bp in length
*    29676    29775: gap of 100 bp
*    29776    30499: contig of 733 bp in length
*    30500    30599: gap of 100 bp
*    30600    31314: contig of 724 bp in length
*    31315    31414: gap of 100 bp
*    31415    32126: contig of 715 bp in length
*    32127    32226: gap of 100 bp
*    32227    32948: contig of 712 bp in length
*    32949    33048: gap of 100 bp
*    33049    33774: contig of 722 bp in length
*    33775    33874: gap of 100 bp
```

```
*  33875 34596: contig of 722 bp in length
*  34597 34696: gap of 100 bp
*  35416 35515: gap of 100 bp
*  35516 36255: contig of 740 bp in length
*  36256 36355: gap of 100 bp
*  36356 37068: contig of 713 bp in length
*  37069 37168: gap of 100 bp
*  37169 37900: contig of 732 bp in length
*  37901 38000: gap of 100 bp
*  38001 38732: contig of 732 bp in length
*  38733 38832: gap of 100 bp
*  38833 39559: contig of 727 bp in length
*  39560 39659: gap of 100 bp
*  39660 40389: contig of 730 bp in length
*  40390 40489: gap of 100 bp
*  40490 41219: contig of 730 bp in length
*  41220 41319: gap of 100 bp
*  41320 42061: contig of 742 bp in length
*  42062 42161: gap of 100 bp
*  42162 42900: contig of 739 bp in length
*  42901 43000: gap of 100 bp
*  43001 43747: contig of 747 bp in length
*  43748 43847: gap of 100 bp
*  43848 44593: contig of 746 bp in length
*  44594 44693: gap of 100 bp
*  44694 45419: contig of 726 bp in length
*  45420 45519: gap of 100 bp
*  45520 46249: contig of 730 bp in length
*  46250 46349: gap of 100 bp
*  46350 47063: contig of 714 bp in length
*  47064 47163: gap of 100 bp
*  47164 47876: contig of 713 bp in length
*  47877 47976: gap of 100 bp
*  47977 48689: contig of 713 bp in length
*  48690 48789: gap of 100 bp
*  48790 49523: contig of 734 bp in length
*  49524 49623: gap of 100 bp
*  49624 50348: contig of 725 bp in length
*  50349 50448: gap of 100 bp
*  50448 51191: contig of 743 bp in length
*  51192 51291: gap of 100 bp
*  51292 52034: contig of 743 bp in length
*  52035 52134: gap of 100 bp
*  52135 52874: contig of 740 bp in length
*  52875 52974: gap of 100 bp
*  52974 53693: contig of 719 bp in length
*  53694 53793: gap of 100 bp
*  53794 54512: contig of 719 bp in length
*  54513 54612: gap of 100 bp
*  54613 55325: contig of 713 bp in length
*  55326 55425: gap of 100 bp
*  55426 56150: contig of 725 bp in length
*  56151 56250: gap of 100 bp
*  56251 56963: contig of 713 bp in length
*  56964 57063: gap of 100 bp
*  57064 57789: contig of 726 bp in length
```

```
Query Match    274;  Conservative    0;  Mismatches   155;  Indels    2;  Gaps    2;
Best Local Similarity   63.6%;   Pred. No. 2e-21;
Matches  274;

Qy    294 gggggtctatagagatggggcaaacta-ccctttcttgttcgggccgtgggctgctact 352
          |  ||||||||| ||||||||||||| |||| ||| |  |||||||||||||||| |||
Db  23029 GGAGCTTATGAGAGTGGGCACACACCCCCTTCCGGTCGGGCGTTGGCTGCTGCTGTATT 22970

Qy    353 tcaaagacggccctc-tttggaaccggtgtgttcggcctccaccctcagtcacccaccgtc 411
          |||    |||||||| || |||| ||||| |||| ||||||||||||||||||||||||||
Db  22969 TCCGCACGCTAANCGTTTGAGATGGTGTCGTGCCCTCGGGCCTCCAGTCGCTCACTGCCCTG 22910

Qy    412 agcgtcggacgcgtacgtggcatcctacacaccgttccgcgcgccaaactgcaggacacccg 471
          |||||| |||||||||| || ||||| |||||||||||||||||||||||||||  ||| |||
Db  22909 AGCGTGGAACGCTATGTGGCCGTGGTGCACCACCACTCCAGGCCAGGTCGTTCCATGGTGACGCGG 22850
```

```
Qy    472 cgccgggccctcaggatcctcggcgcatcgtcttgggcttctcgtgctgctcttctccctgcc 531
          ||  ||||||||||| |||| || ||| |||| ||||||||| | ||||||| ||||||||||
Db  22849 GCCCATGTGCGCGGAGTGCTTGGGGCCGTTGCCGGTTCTTGCCATGCTCTGCTCCCTGCC 22790

Qy    532 aacaccagcatccatggcaaagttccactacttccccaatggggtccctggtcccaggt 591
          |||||||||| |||||| || ||||| ||| ||||||||| ||||||||||| |||||||
Db  22789 AACACCAGCCGTGCAGGCATCCGGGGACTCACGTGCCTCGCGGGGGCCCAGTGCCAGAC 22730

Qy    592 tcggccacctgtacggtcatcaagaccatgtggatctacaaattcatcaccaggtcacc 651
          ||||| ||||| |||||| |||||| ||||||||| ||| |||||||| |||||||||||
Db  22729 TCAGCTGTTTGCATGCGTGGTGTCCGCCCACGGGCCCTCTACAACATGGTAGTGCCAGCAGACCACC 22670

Qy    652 tccttccctattctcacctcctccccatgactgtcatccagtgtcctctactacctcctggca 711
          ||||||| |||||| ||||  |||||||||| ||||||| |||| |||| |||||||||||||
Db  22669 GCGCTGCTCTTCTTCTTGCCTGCCTGCCCATGGCCATCATGAGCGTGCTGCTCTACCTGCTTCATTGGG 22610

Qy    712 ctcagagtgag 722
          || | ||| ||
Db  22609 CTCGGACTGCG 22599
```

```
BASE COUNT      347 a    631 c    598 g    464 t
ORIGIN

Query Match     DIGGDIVGYTEISANVKTMG"
Best Local Similarity   18.5%;   Score 134.8;   DB 88;   Length 2040;
Matches 275;   Conservative 56.1%;   Pred. No. 2.6e-16;
                0;   Mismatches 212;   Indels  3;   Gaps  1;

Qy     128 tccccgtgtctgtggtgtatgtgccaatttttgtggtgggggtcaattggcaatgtcctgg 187
           |  | ||  | |||  | |  |||||   |||||||   |||| ||| || | ||
Db     116 TGCCGGTGACCGCGTGTGTGCCCTGGTGTTCGTGTCGTGCGGGGGTGAGCGGCAAGCTGGTGA 175

Qy     188 tgtgcctggtgattctgcagcaccaggctatgaagagacgcccaccaacttactaccttca 247
           ||  |||||    ||  ||  |||  ||| ||| || ||||| ||||||| ||||||| ||
Db     176 CCGTGAATGCTGAATCGGGGCGGGGCTACCGGGACATGCGGACCACCACCAACTTGTACCTGGGCA 235

Qy     248 gcctggcgggtctctgacctcctggtcctgctccttggaatgccctggaggtctatgaga 307
           |||| ||| | | ||||| || | |||||  ||||||  ||||||||| || | |||
Db     236 GCATTGGCCGCGTTCCGGACCTACTCATCCTTGCTC---GGGCTGCCGTTCGAACCTGTACCGCC 292

Y      308 tgtggtcgcaactacacccttttcttgttcgggcccgtgggctgctacttcaaagacggccctct 367
           | ||| | |  |  ||| ||||  |||||||  |||||  || || || ||| | ||| |
Db     293 TCTGGGCCTCGCGACGCCCTGGGTGTTCGGGCCGCCTGCTCTGCCCGTGTCCCTCTACGTGG 352

Qy     368 ttgagaccgtgtgcttcgcctccatcctcagacatcacacacgtcagcgtggagcgctacg 427
           |  | ||| | |  ||  ||  | ||  | || ||  |||| | ||||| |||| ||||||
Db     353 GCGAGGGCTGCACCTACGCCACGCTGCTGCACATGACCGGCCTCACGGTCCGAGCGGCTACC 412

Qy     428 tggccatcctacaccgttccggcgccaaactgcagagcaacccggggcccctcaagga 487
           ||||  || |  || |||||| | || |  || || |||| || |||||| | |||
Db     413 TGGGCCAATCTGCCCGCGCCGCTCCGCGCCCGGCGGTCTTGGTCACCCGGGCCTTCCCGCGCGC 472

Qy     488 tcctggcatcgtctgggggcttctccgttgctcttctccctgcccaacacacaagcatccatg 547
           ||||  || || |||||| | ||  | ||| ||| |||| || ||| | | ||  |||
Db     473 TCATCGCTGCTGCCGGCGCCTGGGCCTGGCGTTCTCTGCCCGGTCCCCTTCTTGTTCCTGGTGG 532

Qy     548 gcatcaagttccaactacttccccaatgggtccctggtcccaggttcggccacctgtacgg 607
           || || ||  |  ||| ||||| || || | || ||||||||  || | |||| ||||||
Db     533 GCGTCGAGCAGGACCCCGGCATCTCCGTAGTCCCGGGCCTTCAATGGCACGCGCGCGGATCG 592

Qy     608 tcatcaagcc 617
           || ||| |||
Db     593 CCTCCTCGCC 602
```

OM nucleic - nucleic search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: August 30, 2001, 02:02:19 ; Search time 112.8 Seconds
(without alignments)
4057.981 Million cell updates/sec

Title: US-09-684-725-1
Perfect score: 729
Sequence: 1 atggaaaacttcagaatgc..........cactcaagagtgagtatctag 729
Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched: 730101 seqs, 313950809 residues
Total number of hits satisfying chosen parameters: 1460202

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database : N_Geneseq_0601:*
     1: /SIDS8/gcgdata/geneseq/geneseqn/NA1980.DAT:*
     2: /SIDS8/gcgdata/geneseq/geneseqn/NA1981.DAT:*
     3: /SIDS8/gcgdata/geneseq/geneseqn/NA1982.DAT:*
     4: /SIDS8/gcgdata/geneseq/geneseqn/NA1983.DAT:*
     5: /SIDS8/gcgdata/geneseq/geneseqn/NA1984.DAT:*
     6: /SIDS8/gcgdata/geneseq/geneseqn/NA1985.DAT:*
     7: /SIDS8/gcgdata/geneseq/geneseqn/NA1986.DAT:*
     8: /SIDS8/gcgdata/geneseq/geneseqn/NA1987.DAT:*
     9: /SIDS8/gcgdata/geneseq/geneseqn/NA1988.DAT:*
    10: /SIDS8/gcgdata/geneseq/geneseqn/NA1989.DAT:*
    11: /SIDS8/gcgdata/geneseq/geneseqn/NA1990.DAT:*
    12: /SIDS8/gcgdata/geneseq/geneseqn/NA1991.DAT:*
    13: /SIDS8/gcgdata/geneseq/geneseqn/NA1992.DAT:*
    14: /SIDS8/gcgdata/geneseq/geneseqn/NA1993.DAT:*
    15: /SIDS8/gcgdata/geneseq/geneseqn/NA1994.DAT:*
    16: /SIDS8/gcgdata/geneseq/geneseqn/NA1995.DAT:*
    17: /SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT:*
    18: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:*
    19: /SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT:*
    20: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:*
    21: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:*
    22: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	717.8	98.5	1248	1248	21 AAD01123	Human orphan G pro
2	717.8	98.5	1248	1248	21 AAA46022	Human G protein co
3	717.8	98.5	1360	1360	21 AAZ33297	Human neurotensin-
4	280.8	38.5	1209	1209	22 AAF76231	Human G-protein co
5	280.8	38.5	1212	1212	21 AAA30663	Human G protein-co
6	280.8	38.5	1212	1212	21 AAA30739	DNA encoding human
7	279.2	38.3	1212	1212	21 AAZ49707	Human growth hormo
8	243	33.3	1526	1526	21 AAZ49706	Human growth hormo
9	134.8	18.5	1239	1239	21 AAZ45706	Mouse growth hormo
10	134.8	18.5	1390	1390	21 AAZ45404	cDNA encoding the
11	134.8	18.5	2040	2040	21 AAA46116	Human G protein co
12	134.8	18.5	3066		21 AA245402	Genomic sequence o
13	133.2	18.3	1092		20 AA245405	DNA encoding the P
14	127	17.4	1254		17 AA733127	Human neurotensin
15	120.4	16.5	1050		21 AA261492	cDNA encoding canl
16	120.4	16.5	1088		18 AAT68664	Human growth hormo
17	120.4	16.5	1088		18 AAT69756	Human growth hormo
18	120.4	16.5	1101		21 AAA30643	Human G protein-co
19	120.4	16.5	1101		21 AAA30732	Human G protein co
20	120.4	16.5	1101		21 AAZ51463	DNA encoding human
21	120.4	16.5	1122		18 AAT68665	DNA encoding human
22	120.4	16.5	1122		18 AAT69757	Human G protein-co
23	114	15.6	1029		18 AAT68663	Human growth hormo
24	114	15.6	1029		18 AAT69755	Pig growth hormone
25	114	15.6	1063		18 AAT68662	Swine growth hormo
26	114	15.6	1063		18 AAT69754	Pig growth hormone
27	114	15.6	4009		21 AA245967	Swine growth hormo
28	112.4	15.4	1367		21 AA99602	DNA encoding the m
29	112.4	15.4	1368		21 AAT27545	Human D4 receptor
30	112.4	15.4	1370		15 AAQ57655	Recombinant human
31	112.4	15.4	1464		17 AAT27546	Allele D4.2 of the
32	112.4	15.4	1468		17 AAT27547	Recombinant human
33	112.4	15.4	1610		15 AAQ57657	Allele D4.7 of the
34	110.8	15.2	1095		21 AA245993	cDNA encoding the
35	110.8	15.2	1456		15 AAQ57656	Allele D4.4 of the
36	109.2	15.0	1092		18 AAT69760	Rat growth hormone
37	109.2	15.0	1129		18 AAT68667	Rat growth hormone
38	109.2	15.0	3129		18 AAT69759	Rat growth hormone
39	109.4	13.6	1600		19 AAV36462	Rat THR receptor 2
40	99.4	13.4	1551		21 AA249491	Human NT2LP protei
41	97.6	13.4	1575		21 AAV07655	cDNA encoding a hu
42	97.6	13.4	2850		21 AA249492	Human NT2LP protei
43	97.2	13.3	283		19 AAV44930	Human NT2LP protei
44	97.2	13.3	283		19 AAV32651	Galanin receptor G
45	97.2	13.3	283		19 AAV28290	Galanin receptor G

ALIGNMENTS

RESULT   1
AAD01123

ID   AAD01123 standard; cDNA; 1248 BP.
XX
AC   AAD01123;
XX
DT   02-NOV-2000 (first entry)
XX
DE   Human orphan G protein-coupled receptor hRUP6 cDNA.
XX
KW   Human; orphan G protein-coupled receptor; GPCR; hRUP6; drug screening;
KW   transmembrane receptor; signal cascade; ss.
XX
OS   Homo sapiens.
XX
PN   WO20003125B-A2.
XX
PD   02-JUN-2000.
XX
PF   13-OCT-1999; 99WO-US23687.
XX
PR   20-NOV-1998; 98US-0109213.
PR   16-FEB-1999; 99US-0120416.
PR   26-FEB-1999; 99US-0121852.
PR   12-MAR-1999; 99US-0123946.
PR   12-MAR-1999; 99US-0123949.
PR   28-MAY-1999; 99US-0136436.
PR   28-MAY-1999; 99US-0136437.
XX
FH   Key   Location/Qualifiers
FH
FT   CDS   1..1248
FT   /*tag= a
FT   /product= "hRUP6"
FT   /note= "Human orphan G protein-coupled receptor"

```
PR  28-MAY-1999;    90US-0136439.
PR  28-MAY-1999;    90US-0135567.
PR  28-MAY-1999;    90US-0137127.
PR  28-MAY-1999;    90US-0137131.
PR  29-JUN-1999;    90US-0141448.
PR  29-SEP-1999;    90US-0155555.
PR  29-SEP-1999;    90US-0156633.
PR  29-SEP-1999;    90US-0156634.
PR  01-OCT-1999;    90US-0156653.
PR  01-OCT-1999;    90US-0157280.
PR  01-OCT-1999;    90US-0157281.
PR  01-OCT-1999;    90US-0157282.
PR  01-OCT-1999;    90US-0157293.
PR  01-OCT-1999;    90US-0157294.
PR  12-OCT-1999;    90US-0416760.
PR  12-OCT-1999;    90US-0417044.
XX
PA  (AREN-) ARENA PHARM INC.
XX
PI  Chen R,  Dang HT,  Liaw CW,  Lin I;
XX
DR  WPI; 2000-400068/34.
DR  P-PSDB; AAY71296.
XX
PT  Novel human orphan G protein-coupled receptors and the encoding cDNAs
PT  for use in the identification of G protein-coupled receptor agonists -
XX
PS  Claim 21; Page 56-57; 102pp; English.
XX
CC  The present sequence is a cDNA encoding hRUP6, an endogenous human
CC  orphan G protein-coupled receptor (GPCR). The full length hRUP6 cDNA
CC  was cloned by RT-PCR using human thymus cDNA as template.
CC  The orphan GPCR of the invention, like all GPCRs has seven transmembrane
CC  alpha helices with an extracellular N-terminus and an intracellular
CC  C-terminus. However, no endogenous ligands has yet been identified for
CC  the proteins of the invention. The orphan GPCRs may be used in the
CC  identification of their endogenous ligands, and to screen potential GPCR
CC  agonists and antagonists for use as pharmaceutical agents. The proteins
CC  may also be used in the study of GPCR-mediated signalling cascades, and
CC  to elucidate their precise role in normal and diseased human conditions.
CC  Nucleic acid encoding human orphan GPCRs may be used for tissue
CC  localisation expression analysis to provide information about their
CC  function in healthy and pathological states.
XX
SQ  Sequence 1248 BP; 269 A; 393 C; 269 G; 317 T; 0 other;

Query Match             98.5%;  Score 717.8;  DB 21;  Length 1248;
Best Local Similarity   99.7%;
Matches  719;  Conservative    0;  Mismatches   2;  Indels   0;  Gaps    0;

Qy         1 atggaaaaacttcagaaatgcttcctggatctaccaggcagaaactagaagatccattccag  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        10 atggaaaaacttcagaaatgcttcctggatctaccaggcagaaactagaagatccattccag  69

Qy        61 aaacaacctgaacagcaccgaggagtactctggcctcctcctccgacctccgcagccac 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        70 aaacaacctgaacagcaccgaggagtactctggcctcctcctccgacctccgcagccac 129

Qy       121 ttcttcctccccgtgtctgtggtgtatggccaattttgtggtggggtcattggcaat 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       130 ttcttcctccccgtgtctgtggtgtatggccaattttgtggtggggtcattggcaat 189

Qy       181 gtcctggtgtgcctggtgattctgcagcaccaggctatagaagagcgcccaccaactactac 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       190 gtcctggtgtgcctggtgattctgcagcaccaggctatgaagagcgcccaccaactactac 249

Qy       241 ctcttcagcctgcggcgggtctctgacctcctggtctctctccttgaatgccctgaagtc 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       250 ctcttcagcctgcggcgggtctctgacctcctggtctctctccttgaatgccctgaagtc 309

Qy       301 tatgaagatgtggcgcaactaccctttcttgttcgggcccgtggggctgctacttcaaggacg 360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       360 tatgagatgtggcgcaactaccctttcttgttcgggcccgtggggctgctacttcaagacg 360
```

```
Db       310 tatgagatgtggcgcaactaccctttcttgttcgggcccgtggggctgctacttcaagacg 369
Qy       361 gccctctttgagacgtgttgcttcgcctccatcctccagcatcaccacagcgtcagcgtgag 420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 gccctctttgagacgtgttgcttcgcctccatcctccagcatcaccacagcgtcagcgtggag 420
Qy       421 cgctacgtggccatcctacaccgttccgccgcaaactgcaagagaacaccgggccgggcc 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       430 cgctacgtggccatcctacaccgttccgccgcaaactgcaagagaacaccgggccgggcc 489
Qy       481 ctcaggatcctcggggcttctccgtgcctctccctgcccaacacaccagc 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       490 ctcaggatcctcggggcatcgtctgtgggcttctccgtgcctctccctgcccaacacaccagc 549
Qy       541 atccatggcatcaagttccactacttccccaatgggtccctgttcctggtccaggttcggccacc 600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       550 atccatggcatcaagttccactacttccccaatgggtccctggtccaggttcggccacc 609
Qy       601 tgtacggtcatcaagccaccatgtggatctacacaattcatccaggtcacctccttcta 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       610 tgtacggtcatcaagcccatgtggatctacacaattcatccaggtcacctcctta 669
Qy       661 ttctacctcctccccatgactgtcatcagtgtcctactacttcatggcactcagagtg 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1
Db       670 ttctacctcctccccatgactgtcatcagtgtcctactacttcatggcactcagacta 729
Qy       721 a 721
             |
Db       730 a 730
```

```
RESULT  2
AAA46022
AAA46022 standard; cDNA; 1248 BP.
ID   AAA46022 standard; cDNA; 1248 BP.
XX
AC   AAA46022;
XX
DT   22-AUG-2000  (first entry)
XX
DE   Human G protein coupled receptor hRUP6 encoding cDNA SEQ ID NO:11.
XX
KW   Human; G protein coupled receptor; GPCR; transmembrane receptor;
KW   identification; agonist; screening; therapeutic; pharmaceutical;
KW   mutant; ss.
XX
OS   Homo sapiens.
XX
PN   WO20002131-A2.
XX
PD   20-APR-2000.
XX
PF   13-OCT-1999;    99WO-US24065.
XX
PR   13-OCT-1998;    98US-0170496.
PR   12-NOV-1998;    98US-0108029.
PR   20-NOV-1998;    98US-0109213.
PR   27-NOV-1998;    98US-0110060.
PR   16-FEB-1999;    99US-0120416.
PR   26-FEB-1999;    99US-0121852.
PR   12-MAR-1999;    99US-0123944.
PR   12-MAR-1999;    99US-0123945.
PR   12-MAR-1999;    99US-0123946.
PR   12-MAR-1999;    99US-0123948.
PR   12-MAR-1999;    99US-0123949.
PR   12-MAR-1999;    99US-0123951.
PR   28-MAY-1999;    99US-0364436.
PR   28-MAY-1999;    99US-0136437.
PR   28-MAY-1999;    99US-0136439.
PR   28-MAY-1999;    99US-0137127.
PR   28-MAY-1999;    99US-0137131.
PR   28-MAY-1999;    99US-0137567.
PR   30-JUN-1999;    99US-0141448.
PR   27-AUG-1999;    99US-0151114.
```

```
PR   03-SEP-1999; 99US-0152524.
PR   29-SEP-1999; 99US-0156633.
PR   29-SEP-1999; 99US-0155555.
PR   29-SEP-1999; 99US-0156634.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Behan DP, Lehmann-Bruinsma K, Chalmers DT, Chen R, Dang HT;
PI   Gore M, Liaw CW, Lin I, Lowitz K, White C;
XX
DR   WPI; 2000-317986/27.
XX
DR   P-PSDB; AAB02830.
XX
PT   Non-endogenous, human G protein-coupled receptors for screening
PT   receptor, inverse or partial agonists useful as therapeutic agents -
XX
PS   Example 1; Page 85-86; 187pp; English.
XX
CC   The present invention describes transmembrane receptors, preferably
CC   human G protein coupled receptors (GPCR), for which the endogenous
CC   ligand is unknown (orphan GPCR receptors). More specifically the present
CC   invention relates to non-endogenous, constitutively activated versions
CC   of a human GPCR. These non-endogenous human GPCRs can be useful for
CC   the direct identification of candidate compounds as receptors agonists,
CC   inverse agonists or partial agonists for use as pharmaceutical agents.
CC   AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
CC   the exemplification of the present invention.
XX
SQ   Sequence 1248 BP; 269 A; 393 C; 269 G; 317 T; 0 other;

Query Match             98.5%;  Score 717.8;  DB 21;  Length 1248;
Best Local Similarity   99.7%;  Pred. No. 3.3e-146;
Matches  719;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy    1 atggaaaaacttcagaatgcttcctggatctaccacagcagaaaactagaagagatccattccag 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10 atggaaaaacttcagaatgcttcctggatctaccacagcagaaaactagaagatccattccag 69

Qy   61 aaacacctgaacagcaccgaggagtatctggccttctctgcgggcgacctcggcgcagccac 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70 aaacacctgaacagcaccgaggagtatctggccttctctgcgggcgacctcggcgcagccac 129

Qy  121 tcctcctcccccgtgtctgtggtgtatgtgccaatttttgtggtggggtcattggcaat 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  130 tcctcctcccccgtgtctgtggtgtatgtgccaatttttgtggtggggtcattggcaat 189

Qy  181 gtcctggtgtgcctggtgttgattctgcagcgaccaggctatgagaagcgccaccaactactac 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  190 gtcctggtgtgcctggtgttgattctgcagcgaccaggctatgagaagcgccaccaactactac 249

Qy  241 ctcttcaagctgggcgggctctgacctcgctgctcctggaatgccctgggagtc 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  250 ctcttcaagctgggcgggctctgacctcgctgctcctggaatgccctgggagtc 309

Qy  301 tatgagatgtggcgcaactacacctcttcttgttctcggagacg 360
         |||||||||||||||||||||||||||||||||||||||||
Db  310 tatgagatgtggcgcaactacacctcttcttgttctcggagacg 369

Qy  361 gccctcttgagacgtgtgcttcgcctccatcctcagcatcacacaccacgtcagcgtggag 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  370 gccctcttgagacgtgtgcttcgcctccatcctcagcatcacacaccacgtcagcgtggag 429

Qy  421 cgctacgtggccatcctacacccgttccgcgccaaactgcagcgacaccgggcgggcc 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  430 cgctacgtggccatcctacacccgttccgcgccaaactgcagcgacaccgggcgggcc 489

Qy  481 ctcaggatcctcgggggttctccgtgtctctcctgcccaacacacagc 540
         |||||||||||||||||||||||||||||||||||||||||||||
Db  490 ctcaggatcctcgggggttctccgtgtctctcctgcccaacacacagc 549

Qy  541 atccatggcacaagttccactacttccccaatgggtccccaagttcggccacc 600
         ||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db  550 atccatggcatcaagttccactactccccaatgggtccctggtccccaagttcggccacc 609

Qy  601 tgtacggtcatcaagcccatgtggatctacaaattcatccaggtcacctccta 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  610 tgtacggtcatcaagcccatgtggatctacaaattcatccaggtcacctcctta 669

Qy  661 ttctaccttcctcccatgactgtcatcagttctactactcatggcaactcaagagtg 720
         ||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db  670 ttctaccttcctcccatgactgtcatcaggtcctctactacctcatggcaactcaagagacta 729

Qy  721 a 721
         |
Db  730 a 730
```

```
Qy        121 ttcttcctccccgtgtctgtggtgtatgtgccaatttttgtggtgggggtcattggcaat 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 ttcttcctccccgtgtctgtggtgtatgtgccaatttttgtggtgggggtcattggcaat 253

Qy        181 gtcctggtgtgcctggtgattctgcagcaccaggctatgaagacgcccaccaactactac 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 gtcctggtgtgcctggtgattctgcagcaccaggctatgaagacgcccaccaactactac 313

Qy        241 ctcttcagcctggcggtctctgacctcctggtcctgctccttggaatgcccctggaggtc 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 ctcttcagcctggcggtctctgacctcctggtcctgctccttggaatgcccctggaggtc 373

Qy        301 tatgagatgtggcgcaactaccctttcttgttcgggcccgtgggctgctacttcaagacg 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 tatgagatgtggcgcaactaccctttcttgttcgggcccgtgggctgctacttcaagacg 433

Qy        361 gccctctttgagaccgtgtgcttcgcctccatcctcagcatcaccaccgtcagcgtggag 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 gccctctttgagaccgtgtgcttcgcctccatcctcagcatcaccaccgtcagcgtggag 493

Qy        421 cgctacgtggccatcctacacccgttccgcgccaaactgcagagcacccggcgccgggcc 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 cgctacgtggccatcctacacccgttccgcgccaaactgcagagcacccggcgccgggcc 553

Qy        481 ctcaggatcctcggcatcgtctggggcttctccgtgctcttctccctgcccaacaccagc 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        554 ctcaggatcctcggcatcgtctggggcttctccgtgctcttctccctgcccaacaccagc 613

Qy        541 atccatggcatcaagttccactacttccccaatgggtccctggtcccaggttcggccacc 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        614 atccatggcatcaagttccactacttccccaatgggtccctggtcccaggttcggccacc 673

Qy        601 tgtacggtcatcaagcccatgtggatctacaatttcatcatccaggtcacctccttccta 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        674 tgtacggtcatcaagcccatgtggatctacaatttcatcatccaggtcacctccttccta 733

Qy        661 ttctacctcctccccatgactgtcatcagtgtcctctactacctcatggcactcagagtg 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        734 ttctacctcctccccatgactgtcatcagtgtcctctactacctcatggcactcagacta 793

Qy        721 a 721
              |
Db        794 a 794

RESULT 4
AAF76231
ID   AAF76231 standard; cDNA; 1209 BP.
XX
AC   AAF76231;
XX
DT   05-JUN-2001  (first entry)
XX
DE   Human G-protein coupled receptor FM-3 cDNA.
XX
KW   Human; FM-3; G-protein coupled receptor; orphan receptor;
KW   ligand screening; agonist identification; function determination;
KW   signal tranduction disorder; drug discovery; drug design; ss.
XX
OS   Homo sapiens.
XX
PN   WO200114883-A1.
XX
PD   01-MAR-2001.
XX
PF   23-AUG-2000; 2000WO-JP05639.
XX
PR   24-AUG-1999;   99JP-0236597.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Hinuma S,  Hosoya M;
XX
DR   WPI; 2001-218470/22.
XX
PT   Screening compounds to give ligands of orphan receptor protein or its
PT   subtype, antagonists and agonists with efficiency and reliability as
PT   drugs for treating various diseases e.g. abnormality in intracellular
PT   signal transduction  -
XX
PS   Examples; Page 66-67; 70pp; Japanese.
XX
CC   The invention relates to a method for screening for ligands of an orphan
CC   receptor protein to identify compounds that can promote or inhibit its
CC   function. The method comprises contacting a potential ligand with cells
CC   which express the orphan receptor of interest, and measuring the degree
CC   of receptor-mediated cell stimulatory activity for the test compound.
CC   This measurement may then be compared with that for a known control, and
CC   an agonist selected. The structures of all the compounds tested,
CC   particularly the chosen agonists, may then be compared to identify
CC   ligand features that correlate with agonist activity. This information
CC   can be used to predict ligand candidates in order to provide an assay of
CC   the specific binding capability to pinpoint receptor activity. The
CC   invention also encompasses the compounds identified using the method of
CC   the invention. The method can be used to screen for ligands of an
CC   orphan receptor or subtype thereof, and thereby to efficiently and
CC   reliably identify antagonists and highly active agonists. The compounds
CC   identified may be used as drugs for treating a variety of diseases,
CC   particularly those arising from an abnormality of intracellular signal
CC   transduction. Such drugs are likely to be safe and have a low toxicity.
CC   The exemplifications describe the process of screening for ligands
CC   of the human G protein-coupled receptor FM-3. The candidate compounds
CC   are the peptides given in records AAB73357-AAB73375, and the identified
CC   ligands are H-2980 (AAB73358), F-8-F-NH2 (AAB73361) and small
CC   cardioactive peptide B (SCPB; AAB73375). The present sequence represents
CC   cDNA encoding human FM-3 which was cloned into a mammalian expression
CC   vector for expression in CHO (Chinese hamster ovary) cells.
XX
SQ   Sequence 1209 BP; 192 A; 405 C; 366 G; 246 T; 0 other;

  Query Match             38.5%;  Score 280.8;  DB 22;  Length 1209;
  Best Local Similarity   65.8%;  Pred. No. 6.4e-52;
  Matches  408;  Conservative    0;  Mismatches  212;  Indels    0;  Gaps    0;

Qy        103 ggacctcggcgcagccacttcttcctccccgtgtctgtggtgtatgtgccaatttttgtg 162
              || || | ||  |   |  | ||| | ||| | | ||     ||  |||  || || |||
Db         88 gggccccagcagacagagctgttcatgcccatctgtgccacatacctgctgatcttcgtg 147

Qy        163 gtgggggtcattggcaatgtcctggtgtgcctggtgattctgcagcaccaggctatgaag 222
              ||||| |   | ||||||| |||   || ||||| || ||||  ||| ||| |||   
Db        148 gtgggcgctgtgggcaatgggctgacctgtctggtcatcctgcgccacaaggccatgcgc 207

Qy        223 acgcccaccaactactacctcttcagcctggcggtctctgacctcctggtcctgctcctt 282
              ||||| |||||||||||||||||||||||||| || || ||||| ||||| |||||  | 
Db        208 acgcctaccaactactacctcttcagcctggccgtgtcggacctgctggtgctgctggtg 267

Qy        283 ggaatgcccctggaggtctatgagatgtggcgcaactaccctttcttgttcgggcccgtg 342
              ||  ||||||||||| |||||||||||||||| |||||||||| ||| || | ||    |  
Db        268 ggcctgcccctggagctctatgagatgtggcacaactaccccttcctgctgggcgttggt 327

Qy        343 ggctgctacttcaagacggccctctttgagaccgtgtgcttcgcctccatcctcagcatc 402
              |||||||| |||   |||   || ||||||| |  || ||| | |||||  | |||| | ||
Db        328 ggctgctatttccgcacgctactgtttgagatggtctgcctggcctcagtgctcaacgtc 387

Qy        403 accaccgtcagcgtggagcgctacgtggccatcctacacccgttccgcgccaaactgcag 462
              ||  || | |||||||| ||||| |||||| |  | |||||  |||  ||||       |
Db        388 actgccctgagcgtggaacgctatgtggccgtggtgcacccactccaggccaggtccatg 447

Qy        463 agcacccggcgccgggccctcaggatcctcggcatcgtctggggcttctcccgtgctcttc 522
                 || |||  ||  |  | | |  | || ||   |||||||||  |  || |||||| |
Db        448 gtgacgcgggcccatgtgcgccgagtgcttggggccgtctggggtcttgccatgctctgc 507

Qy        523 tccctgcccaacaccagcatccatggcatcaagttccactacttccccaatgggtccctg 582
              ||||||||||||||||| | ||  ||||||   |    |    || | |||     ||  ||
```

```
Db    508 tccctgccaacaccagcctgcacgtcacgtgcacgtgccctgcgggcca 567
          ||||  |  |  |  |||    ||||||   |||||||||   |||||
Qy    583 gtcccaggttcggccacctgtacgtcatcaagccatgtggatcacaatttcatatc 642

Db    568 gtgcaggactcagctgcgttgcatcgtggtcgcgccacaggtcctctacac 627

Qy    643 caggtcacctccttcattctcacctcccccatgactgtcatcagtgtcctctactac 702

Db    628 cagaccaccgcgtcctcttctctgcctccgccatgctggcca 687

Qy    703 ctcatggcactcaagagtgag 722
          |||| |   ||    |   |
Db    688 ctcatgggctgcgactgcg 707
```

```
RESULT  5
ID   AAA30663 standard; cDNA; 1212 BP.
XX
AC   AAA30663;
XX
DT   21-AUG-2000 (first entry)
XX
DE   Human G protein-coupled receptor MIG cDNA.
XX
KW   G protein-coupled receptor; GPCR; constitutively active;
KW   intracellular loop 3; transmembrane domain 6; drug screening;
KW   agonist; antagonist; ss.
XX
OS   Homo.sapiens.
XX
PN   WO20022129-A1.
XX
PD   20-APR-2000.
XX
PF   12-OCT-1999;   99WO-US23938.
PR   13-OCT-1998;   98US-0170496.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Behan DP, Chalmers DT, Liaw CW;
XX
DR   WPI; 2000-329165/28.
DR   P-PSDB; AAY90638.
XX
PT   Non-endogenous constitutively activated human G protein-coupled
PT   receptors, useful for identifying agonists for use as pharmaceutical
PT   agents
XX
PS   Example 1; Page 188-189; 341pp; English.
XX

CC   The invention relates to constitutively active, non-endogenous versions
CC   of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-
CC   AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743
CC   and AAA30775-A3079). The mutant proteins of the invention contain a
CC   mutation in a portion of the protein comprising intracellular loop 3
CC   (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,
CC   is substituted for an endogenous residue in IC3 at a position 16 amino
CC   acids N-terminal of an endogenous proline in TM6 to form a sequence
CC   X-(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg
CC   or Ala, and is preferably Lys. When the endogenous residue at this
CC   position is replaced by His, Arg or preferably Ala.
CC   The 15 amino acid stretch between the substituted amino acid and the Pro
CC   may be endogenous, non-endogenous, or a mixture of endogenous and
CC   non-endogenous residues. The constitutively active GPCRs are useful for
CC   identifying antagonists, agonists and partial agonists for use as
CC   pharmaceutical agents. The mutant proteins are also useful in research
CC   settings for elucidating the roles of the receptors in normal and
CC   diseased conditions. Antagonists for a particular GPCR are useful for
CC   treating diseases and disorders associated with that receptor. Because
CC   the novel mutant GPCRs are constitutively active, they can be used
CC   directly for screening of compounds without the need for endogenous
```

```
CC   ligands. The present sequence represents cDNA encoding a human wild-type
CC   GPCR used in an exemplification of the invention. This was cloned and
CC   subjected to site-directed mutagenesis (SDM) to generate DNA encoding
CC   the corresponding mutant of the invention.
XX
SQ   Sequence 1212 BP; 193 A; 405 C; 367 G; 247 T; 0 other;

Query Match             38.5%;  Score 280.8; DB 21; Length 1212;
Best Local Similarity   65.8%;  Pred. No. 6.4e-52;
Matches  408; Conservative    0; Mismatches 212; Indels    0; Gaps   0;

Qy    103 ggacctccgcgcagcaaccacactcttcctccccgtgtctgtggtctgtgtgtccaatttgtg 162
          |||         ||    |||   |||||||||||||||||||   |||||     |||||
Db     88 gggcccaagcgcagacagaagctgttcatgcccatctgtcacaacatacctgtgatctcgtg 147

Qy    163 gtggggtcattggcaaatgtccctggtttgcctggtgattctgcagcacaggctatgaag 222
Db    148 gtggggcgtcgtgggcaatgggcatctggtcatcctgcgcccacaaggccatgcgc 207

Qy    223 acgcccaccacaactactactcttcctcaagcctggctgttctgaccctggttctgcctt 282
Db    208 acgcctaccacaactactactcttcctcagcctggctgttctgaccctgctggtg 267

Qy    283 ggaatgccctgaggtctcatgtggcgcaactacctttcttgttcgggccgtg 342
Db    268 ggcctgccctgggagctctcatgagatgtggcacaactacctcctgcgggcgttggt 327

Qy    343 ggctgctacttcaagacggccctcttcttgagacgtgtgtctttcctccatccctcagcatc 402
Db    328 ggctgctatttccgccacactctgcttgtttgagatggtctctgcctcagtgctcaagtc 387

Qy    403 accacgtcaagcgttggagcggctacgtggcatctcactacaccccgtttccgcgccaaactgcag 462
Db    388 actgccctgagcgttggaaacgctatgtcggcgtctgcaccccactccaggcaggtccatg 447

Qy    463 agcaacccgcgtcgcgggtccctcaggaatcctcggcatcgtctcgggggttctctgcttcttc 522
Db    448 gtgacgcggggcccatgtcgccgagtgcttctgggggtcttgccatgctctgc 507

Qy    523 tccctgccaacaccagcctgcacgtcacgtgcacgtgccctgcgggccccaatggtccctg 582
Db    508 tccctgccaacaccagcctgcacgtcacgtgcacgtgccctgcgggcca 567

Qy    583 gtcccaggttcggccacctgtacgtcatcaagccatgtggatcacaatttcatatc 642
Db    568 gtgcaggactcagctgcgttgcatcgtggtcgcgccacaggtcctctacac 627

Qy    643 caggtcacctccttcattctcacctcccccatgactgtcatcagtgtcctctactac 702
Db    628 cagaccaccgcgtcctcttctctgcctccgccatgctggcca 687

Qy    703 ctcatggcactcaagagtgag 722
Db    688 ctcatgggctgcgactgcg 707
```

```
RESULT  6
ID   AAA30739 standard; DNA; 1212 BP.
XX
AC   AAA30739;
XX
DT   21-AUG-2000 (first entry)
XX
DE   DNA encoding human mutant G protein-coupled receptor MIG (T273K).
XX
KW   G protein-coupled receptor; GPCR; constitutively active;
KW   intracellular loop 3; transmembrane domain 6; drug screening;
KW   agonist; antagonist; mutant; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
```

XX
PN   WO20002129-A1.
XX
PD   20-APR-2000.
XX
PF   12-OCT-1999;   99WO-US23938.
XX
PR   13-OCT-1998;   98US-0170496.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Behan DP, Chalmers DT, Liaw CW; .
XX
DR   WPI; 2000-329165/28.
DR   P-PSDB; AAY90673.
XX
PT   Non-endogenous constitutively activated human G protein-coupled
PT   receptors, useful for identifying agonists for use as pharmaceutical
PT   agents -
XX
PS   Example 2; Page 288; 34lpp; English.

CC   The invention relates to constitutively active, non-endogenous versions
CC   of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-
CC   AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743
CC   and AAA30775-A30779). The mutant proteins of the invention contain a
CC   mutation in a portion of the protein comprising intracellular loop 3
CC   (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,
CC   is substituted for an endogenous residue in IC3 at a position 16 amino
CC   acids N-terminal of an endogenous proline in TM6 to form a sequence
CC   X-(AA) 15-Pro. The endogenous amino acid is selected from Lys, His, Arg
CC   or Ala, and is preferably Lys. When the endogenous residue at this
CC   position is Lys, this residue is replaced by His, Arg or preferably Ala.
CC   The 15 amino acid stretch between the substituted amino acid and the Pro
CC   may be endogenous, non-endogenous, or a mixture of endogenous and
CC   non-endogenous residues. The constitutively active GPCRs are useful for
CC   identifying antagonists, agonists and partial agonists for use as
CC   pharmaceutical agents. The mutant proteins are also useful in research
CC   settings for elucidating the roles of the receptors in normal and
CC   diseased conditions. Antagonists for a particular GPCR are useful for
CC   treating diseases and disorders associated with that receptor. Because
CC   the novel mutant GPCRs are constitutively active, they can be used
CC   directly for screening of compounds without the need for endogenous
CC   ligands. Sequences AAA30709- AAA30743 and AAA30775-A30779 represent DNAs
CC   encoding the mutant human GPCRs of the invention.
XX
SQ   Sequence 1212 BP; 194 A; 403 C; 368 G; 247 T; 0 other;

Query Match   38.5%; Score 280.8; DB 21; Length 1212;
Best Local Similarity   65.8%; Pred. No. 6.4e-52;
Matches 408; Conservative   0; Mismatches 212; Indels   0; Gaps   0;

Qy   103 ggacctcggcgcagccacttcttcctcccccgtgtctgtggtgtatgtgcaattttgtg 162
          ||||| ||| | || ||||| | | ||||||||| || | ||| |||| |||||
Db   88 gggcccagagaagcagagctgttcatcgcacataccctgttgcaatttctc 147

Qy   163 gtggggtcattgcaaatgtcctggtgtgcctgtggtgttcgcaagcaccaggctatgaag 222
          |||||| || | ||| | | ||||| ||| |||| | ||| || || ||
Db   148 gtggggcgctgtgggcaatgggctgacctgtctgtcatcctgcgcccacaaggccatggcg 207

Qy   223 acgcccacacaactactactctcttcaagcctgacctccggtgtctcctt 282
          ||||| | | || | || | ||| | ||| ||| || || |||||||
Db   208 acgcctaccaactactactctcttcaagcctgctcggacctggctgctccgg 267

Qy   283 ggaatgcccctggaggtcatcatgagatggcgcaactaaccctttcttgttcgggccgtg 342
          ||| || | ||| || || | | | ||| | || |||| | ||| || |||
Db   268 ggcctgccccctggacctcatgagatgtgcaaactacccctctggtggt 327

Qy   343 ggctgctacttcaagagaccccggccctccatcctcagcatc 402
          ||||| || | || || || || || || | ||| |||
Db   328 ggctgctattccgcacgctgtttggagatggtctggcgttggt 387

Qy   403 accaccgtcagcgtggagcgctacgtggccatctacaaccgttccgcgccaaactgcag 462
          || || | |||||||| || || || ||| |||||| || || || | ||||
Db   388 actgccctgagcgtggaacgctatgtggccgctggtgaccccactccaggccagttccatg 447

Qy   463 agcaccggcgcccggccgggcccctcaggatcctcggcatgtctcggggcttctccgtgtcttc 522
          || || | | || | | || | | || ||| || | ||| || ||| |
Db   448 gtgacgcggccccatgtgcggccgagtgttggggccgtctgggtctggcatgtgctgc 507

Qy   523 tccctgcccaaccaccagatccatggcatcaagttccactactcccaatgggtccctg 582
          | ||||||| || | | | ||| ||||| || ||| | ||| ||| ||||||
Db   508 tccctgcccaaccacccgaccagctgaccggcatccgcagtcgcggcctgcccctggtg 567

Qy   583 gtcccaggttcggccacctgtacggtcatcaagcccacgtggatctacaaatttcatcatc 642
          || || | || | |||| | ||| ||| | || | ||| ||| | ||| ||
Db   568 gtgccagactgagctgtttgcatgtccgaggggccctcacaaaatggtagtg 627

Qy   643 caggtcaccttcttcctattctactcccactctccctccgtcccctcagtgtctccatcactac 702
          |||| ||||| ||| | | ||||||||| | | || | | ||| || || | | ||| |
Db   628 cagaccaccgcgtgctcttcttctgctcccacatggccacatgagcgcgtctcttacctg 687

Qy   703 ctcatggccatcaagagtgag 722
          ||||| || | || | |
Db   688 ctccattgggcgcgactcgg 707

RESULT 7
AC   AA249707
ID   AA249707 standard; DNA; 1212 BP.
XX
AC   AAA249707;
XX
DT   18-APR-2000 .(first entry)
XX
DE   Human growth hormone secretagogue related receptor DNA.
XX
KW   Human growth hormone secretagogue related receptor; GHSR-R;
KW   mouse GHSR-R gene; ligand; agonist; antagonist; GHSR-R agonist;
KW   neuromedin U; obesity; depression; anxiety; compulsion; neurose;
KW   insomnia; sleep disorder; substance abuse; pain; neuroprotective;
KW   cognitive disorder; memory enhancement; Alzheimer's disease;
KW   GHSR-R antagonist; anorectic; nootropic; antidepressant; tranquilizer;
KW   antiaddictive; analgesic; ds.
XX
OS   Homo sapiens.
XX
FH
FT   Key   Location/Qualifiers
FT   CDS   1..1212
FT   /tag= a
FT   /product= a
FT   receptor= "Human growth hormone secretagogue related
FT   /transl_except= (pos: 646..648, aa: Phe)
FT   /transl_except= (pos: 748..750, aa: Gly)
FT   /transl_except= (pos: 858..900, aa: Val)
XX
PN   WO20002919-A1.
XX
PD   20-JAN-2000.
XX
PF   13-JUL-1999;   99WO-US15941.
XX
PR   13-JUL-1998;   98US-0092623.
XX
PA   (MERI) MERCK & CO INC.
XX
PI   Liu Q,   Howard AD,   Mckee KK;
XX
DR   WPI; 2000-147599/13.
DR   P-PSDB; AAY44642.
XX
PT   New receptors useful for identifying agonist and antagonist for
PT   treating or preventing obesity, anxiety, neuroses, insomnia -
XX
PS   Claim 10; Fig 2; 50pp; English.

```
XX
CC  The present sequence encodes human growth hormone secretagogue related
CC  receptor (GHSR-R). This was isolated from a human PAC library and
CC  sequenced using dye terminator cycle sequencing ready reactions. This
CC  exhibits 63% identity with mouse GHSR-R gene. GHSR-R is useful for
CC  identifying ligands, agonist or antagonist which binds to it. A GHSR-R
CC  agonist or neuromedin U is used in the treatment of obesity, depression,
CC  anxiety, compulsion, neuroses, insomnia/sleep disorders, substance abuse,
CC  pain, neuroprotective and cognitive disorders and memory enhancement
CC  including Alzheimer's disease. GHSR-R antagonist is useful for increasing
CC  food intake of a mammal.
XX
SQ  Sequence 1212 BP; 193 A; 406 C; 366 G; 247 T; 0 other;

Query Match    38.3%;  Score 279.2;  DB 21;  Length 1212;
Best Local Similarity    65.6%;
Matches  407;  Conservative    0;  Pred. No. 1.4e-51;  Indels    0;  Gaps    0;

Qy  103  ggacctcggcgcagccacttcttcctcccccgtgtctgtgttgtatgtgccaattttgtg 162
         ||||||  ||  ||||| |||||     |||||     |||    ||| |||   |||||||||
Db   88  gggcccagcagacgagcgttcatcgtcatgcccatctgtgcccacatactgctgtccgtg 147

Qy  163  gtggggtcattggcaatgtcctggtgtgtgctggtggaccaccaagccagctatgaag 222
         ||||| ||  | |||||||  |||| ||||| ||||    ||   ||||||||||| |||
Db  148  gtgggcgctgtgggcaatggcctgaactgtctcggtcatcctcgccacaagccatgcgc 207

Qy  223  acgcccaccaactactactctttcagcctggttgtctcctggtctgctcctt 282
         |||||  |||||||||||| ||||||||| | |||| |||| ||||||  ||
Db  208  acgcctaccaactactacctcttcagcctgccgtcggttcatcctcgcctgtg 267

Qy  283  ggaatgcccctggaggtctatgagatgtggcgcaacacacccctttcttgttcgggcccgtg 342
         ||  || |||||||| ||||| |||||||||| |||||||| |||||||| |||| | ||| |
Db  268  ggcctgccctggagctcatgtagagatgtgcacaacaactacccctttctgttcgcgttggt 327

Qy  343  ggctgctacttcaagacggccctctcttttgagaccgtgtgtcctccatcctcagcgatc 402
         |||||||| |   | ||| |||| |||| ||||| ||| ||  | ||| ||||| ||| |||
Db  328  ggctgctattccgcacgctactgtttgagatgttctgcctggcctcagtgctcaaagtc 387

Qy  403  accaccgtcagcgtggagcgtgacggcaggcatcctccacaccgttccgcgcaaactgcag 462
         |  |||| |||||||||| ||| |  ||||| |||  |||||| || || |   ||| ||
Db  388  actgcctgagcgtggaacgctatgtggccgctgtgacccactccaggcaggtccatg 447

Qy  463  agcaccccggcgccgggcccctcaggaatcctcggggcttctccgtgctcttc 522
         || || |||||||| |||  | || |||||  | |||| ||||||||||  ||
Db  448  gtgacgcggcgcccatgtcgccgagtgcttggggccgtctgggctcttgcatgtctgc 507

Qy  523  tccctgcccaacacagccatccaagttcaactacttcccaatgggtccctg 582
         |||||||| ||||| ||| | ||| |||||||| ||||| |||||| ||||
Db  508  tccctgccaacaccagcctgtacaacatcgacgctgcacgtgccctgcccggggccca 567

Qy  583  gtccaggttggccacctgtacgtgcatccaagccatgtggattcacaaattcatcatc 642
         ||||| || |||||||||||||| |||||| |||| |||| || ||||| ||||||||
Db  568  gtgccagactcagctgttgcatgtggtccccatctacaacatggtagtg 627

Qy  643  caggtcacctccttcctcctcccatgactgtcatcagtgtcctctactac 702
         ||  ||||| |||| |||||||| |||| |||||||||| |||||| |||
Db  628  cagacaccggcgtcttgttctctgccacagtcatggagcgtggctaccctg 687

Qy  703  ctcatggcaactcaagagtgag 722
         ||||| |  ||||||||  ||
Db  688  ctcattgggctgcgactgcg 707
```

```
DE  Mouse growth hormone secretagogue related receptor DNA.
XX
DT  18-APR-2000  (first entry)
XX
AC  AAZ49706;
XX
ID  AAZ49706 standard; DNA; 1526 BP.
XX
RESULT   8
AAZ49706
```

```
XX
KW  Mouse growth hormone secretagogue related receptor; GHSR-R;
KW  human GHSR-R gene; ligand; agonist; antagonist; GHSR-R agonist;
KW  neuromedin U; obesity; depression; anxiety; compulsion; neuroses;
KW  insomnia; sleep disorder; substance abuse; pain; neuroprotective;
KW  cognitive disorder; memory enhancement; Alzheimer's disease;
KW  GHSR-R antagonist; anorectic; nootropic; antidepressant; tranquilizer;
KW  antiaddictive; analgesic; ds.
XX
OS  Mus sp.
XX
FH  Key            Location/Qualifiers
FT  CDS            309..1526
FT                 /*tag= a
FT                 /product= "Mouse growth hormone secretagogue related
FT                 receptor"
XX
PN  WO200002919-A1.
XX
PD  20-JAN-2000.
XX
PF  13-JUL-1999;  99WO-US15941.
XX
PR  13-JUL-1998;  98US-0092623.
XX
PA  (MERI ) MERCK & CO INC.
XX
PI  Liu Q, Howard AD, Mckee KK;
XX
DR  WPI; 2000-147599/13.
DR  P-PSDB; AAY44641.
XX
PT  New receptors useful for identifying agonist and antagonist for
PT  treating or preventing obesity, anxiety, neuroses, insomnia -
XX
PS  Claim 9; Fig 1; 50pp; English.
XX
CC  The present sequence encodes mouse growth hormone secretagogue related
CC  receptor (GHSR-R). This was isolated from a mouse EST derived from a
CC  T-cell library. This exhibits 63% identity with human GHSR-R gene. GHSR-R
CC  is useful for identifying ligands, agonist or antagonist which binds to
CC  it. A GHSR-R agonist or neuromedin U is used in the treatment of
CC  obesity, depression, anxiety, compulsion, neuroses, insomnia/sleep
CC  disorders, substance abuse, pain, neuroprotective and cognitive disorders
CC  and memory enhancement including Alzheimer's disease. GHSR-R antagonist
CC  is useful for increasing food intake of a mammal.
XX
SQ  Sequence 1526 BP; 301 A; 453 C; 395 G; 377 T; 0 other;

Query Match    33.3%;  Score 243;  DB 21;  Length 1526;
Best Local Similarity    61.7%;
Matches  387;  Conservative    0;  Pred. No. 9.6e-44;  Indels    0;  Gaps    0;

Qy  103  ggacctcggcgcagccacttcttcctcccccgtgtctggtgttgtatgtgccaattttgtg 162
         ||||| || ||  || ||||| ||||||||  | |||| || || |||| ||||||||||||
Db  396  gggccacagcagagtgaaacagtttgtccccatcctgtgtcaccgtgatcttccggctgtg 455

Qy  163  gtggggtcattggcaatgtcctggtgtgtgctggtggaccaccaagccagctatgaag 222
         ||||| ||  | |||||||  |||| ||||| ||||    |||  ||||||  | |||
Db  456  gtgggcactgtgggcaaggggctgaccttgtcctgtgcagacaccaggctatgagg 515

Qy  223  acgcccaccaactactactctttcagcctggttgtctcctggtcacgtggcaaagagc 282
         ||||||||||||||||||| |||||| || | ||| ||||| |  ||||   || ||
Db  516  acgcccaccaacttctacctccagcctcgtcgatatgctggtcgccgatatgtcctg 575

Qy  283  ggaatgcccctgaggtctatgagatggtgcgcaactacccttttctgtttcggcccctg 342
         ||||||  |||||| |||||||||||  || ||||||| | |||| ||| |||||  ||
Db  576  ggcttgcctctggagcttatggagatggcagcaaaattaccgttccagcttcctttggg 635

Qy  343  ggctgctacttcaagacggccctctcttttgagacctgtgttcttcggcctccatcagcatc 402
         |||||||| |||||| || |||| |||||||| | || ||| |||| ||||| |||||||||
Db  636  gcctgctacttccgaatctgtcttagagaccgtctgcctagcttcagtgctcaaagtc 695
```

```
Qy    403 accaccgtcagcgtcggagcgctaagtgggccatcctacacaccgttccgcgccaaactgcag 462
          ||  || |||||||  |  || ||  ||||| |||  ||  ||  |||
Db    696 acagcccggagtgtcggagcgttatcgtggccgtgtgtcgccccactccaagccaagtctgtg 755

Qy    463 agcacccgcgccgggccctcaaggatcctcggcgatcgtctggggcttctccgtgctcctc 522
          |  ||  |||||||| ||  ||  || |||  ||  || ||
Db    756 atgacaacgggccccatcgtgcgccgcatggttcggggcgtcctcgctactctcttc 815

Qy    523 tccctgccacaccagccatccatgcatcaagttccactacttccccaatgggtccctg 582
          || ||||| || || ||||||  ||  ||||  ||||||  ||| |  |||  |||
Db    816 tctctgcccaaacccgtgcatggcctcagtcaactaactgtgcctgccacagtccctg 875

Qy    583 gtcccaggttcggccacctgtacgtcatcaagccacatggttggatctacaaatttcaatc 642
          ||  ||  ||  || || |||  ||  || || || |||| || || ||||   | ||||
Db    876 gtgcccgactcagctatatgttcgctggtgggtgccccatggacaccttctacaagttcggggta 935

Qy    643 caggtcacctccttcttctattctaacctctccatcagtcagtcgtcctcctactac 702
          |||  | || |  || ||  || || |||  ||| |||| || || |||  ||| ||
Db    936 ctggactaccgcaactgtctcttcttctgtcccagtgtcacatggttgctgtatctg 995

Qy    703 ctcatggcacttcaagtgtgagtatctag 729
          || ||||  |   || ||
Db    996 cttcattggggctgcggctcggagggag 1022
```

```
PS   Claim 4; Fig 2; 44pp; English.
XX
CC   The present sequence encodes splice variant MTL-R1A of the motilin
CC   receptor. The gene encodes a G-protein coupled receptor, and is
CC   designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,
CC   MTL-R1A (see AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a
CC   functional seven transmembrane domain form, and MTL-R1B is a truncated
CC   five transmembrane domain. The MTL-R1 proteins are used to identify
CC   agonists and antagonists which can be used for treating gastric motility
CC   disorders, functional defects, disorders secondary to neurological
CC   disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
CC   dysmotility, diabetes, infections, stress-related motility disorders,
CC   psychgenic disorders, gastroparesis, gastro-oesophageal reflux disease,
CC   constipation, chronic idiopathis pseudo obstruction, acute faecal
CC   impaction, postoperative ileus, gallstones, infantile collic, irritable
CC   bowel syndrome, non-ulcer dyspepsion, non-cardiac chest pain and
CC   diarrhoea. They can also be used in the preparation for colonoscopy,
CC   endoscopy and duodenal intubation.
XX
SQ   Sequence 1239 BP; 172 A; 431 C; 397 G; 239 T; 0 other;

Query Match            18.5%;  Score 134.8;  DB 21;  Length 1239;
Best Local Similarity   56.1%;
Matches  275;  Conservative  0;  Mismatches 212;  Indels  3;  Gaps  1;
```

```
Qy    128 tccccgtgtctgtggttgtatgtgccaagatcttttgtgtggggtgtcattggcaaatgtcctg 187
          |  || |||| |  | | || ||| ||  || || | |   ||    ||   |||  || ||
Db    116 tgccggtgaccgctgtgtgtcgcctgtgcctgtctcgtcgtggggtgagcggcaacgtggtga 175

Qy    188 tgtgcggtgattctgcacacacaggcgtatgtgaagacgcccacacaacttctacctca 247
          |  || |||| ||| ||   || |  ||  |  |  ||||| ||||| ||| ||| ||
Db    176 ccgtgatgctgatcgggcgcgctacccgggacatgggaccacacaacttgtacctgggca 235

Qy    248 gccttggccggtctctgacctcctggtcctgcttcctggaatgcccccctgggaggtctatgaga 307
          || ||||| ||||| ||| |||| |  || || || ||   ||   |||||||||| || ||||
Db    236 gcaatggccgtgtccgacatcactcctctc---gggctgccgttcgacctgtaccgcc 292

Qy    308 tgtgtgcgcaactacccttcttcttcgttccggggccgtgggctgctacttcaagaacggccct 367
          |  || |||||| || | || ||  || ||  | ||| || ||| ||  || ||| |||||
Db    293 tctggcgcttcgggccctggttcggggccgcgtcctgtccgcgcctgtcccctacctgg 352

Qy    368 ttgagacggtgtgcttcgcctccatcctcaagcacacccacccgtctcctgaagcgtgggagcgctacg 427
          || || |||| || ||  ||  |||| |||||||||| |||  | |  | |   || || ||| ||||
Db    353 gcgaggggctgcacctacgccacgctgctgcacatgaccgcgctcgagcgctac 412

Qy    428 tggccatcctacaccccgtccgcgccgccaaactgcagacagcacccggcgcccggcgggccctcagga 487
          |||||||  |  | ||||| || ||  | || |  |||  |||| |  ||||| ||| ||||| ||
Db    413 tggccatctgccccgccctccggcccttcttggtcaccgcgccgtccgcgcgc 472

Qy    488 tcctccggcatcgtctgggggcttctccgtgctcctcaccggccccccaacacagcaatccatg 547
          ||||| ||||| || ||||| || ||| | || || |||||  ||| ||||| ||||| ||||
Db    473 tcatcgtgtgctgctggggctgctctcgccggtccccttcttcctgttcctggtgg 532

Qy    548 gcatcaagttccactacttccccaatgggtccctggttcggccactgtacg 607
          || || |||  | || ||||| |||||| ||| |||  |||  ||||||||
Db    533 gcgtcgacgcaggagaccccgcatccgtagtccggcctcaatggcaccgccggatcg 592

Qy    608 tcatcaagc 617
          ||||| |||
Db    593 cctcctcgcc 602
```

```
RESULT  10
AA245404
ID   AA245404 standard; cDNA; 1390 BP.
XX
AC   AA245404;
XX
DT   27-MAR-2000 (first entry)
XX
DE   cDNA encoding the motilin receptor splice variant MTL-R1B.
XX
```

```
RESULT  9
AA245403
ID   AA245403 standard; cDNA; 1239 BP.
XX
AC   AA245403;
XX
DT   27-MAR-2000 (first entry)
XX
DE   cDNA encoding the motilin receptor splice variant MTL-R1A.
XX
KW   Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
KW   spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
KW   functional defect; neurological disorder; scleroderma; colonoscopy;
KW   paraneoplastic syndrome; radiation induced dysmotility; diabetes;
KW   infection; stress-related motility disorder; psychgenic disorder;
KW   gastroparesis; gastro-oesophageal reflux disease; constipation;
KW   chronic idiopathis pseudo obstruction; acute faecal impaction;
KW   postoperative ileus; gallstones; infantile collic; diarrhoea;
KW   irritable bowel syndrome; non-ulcer dyspepsion; non-cardiac chest pain;
KW   endoscopy; duodenal intubation; ds.
XX
OS   Homo sapiens.
XX
FH   Key              Location/Qualifiers
FT   CDS              1..1239
FT                    /*tag= a
FT                    /product= "MTL-R1A"
XX
PN   WO9964436-A1.
XX
PD   16-DEC-1999.
XX
DT   16-DEC-1999.
XX
PF   08-JUN-1999; 99WO-US12773.
XX
PR   12-JUN-1998; 98US-0089098.
XX
PA   (MERI ) MERCK & CO INC.
XX
PI   Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
PI   Pong S, Smith RG;
XX
DR   WPI; 2000-105868/09.
DR   P-PSDB; AAY54145.
XX
PT   Novel receptor protein for screening compounds used in treating
PT   irritable bowel syndrome, constipation and other gastric conditions -
```

KW Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
KW spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
KW functional defect; neurological disorder; scleroderma; colonoscopy;
KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;
KW infection; stress-related motility disorder; psychgenic disorder;
KW gastroparesis; gastro-oesophageal reflux disease; constipation;
KW chronic idiopathis pseudo obstruction; acute faecal impaction;
KW postoperative ileus; gallstones; infantile collic; diarrhoea;
KW irritable bowel syndrome; non-ulcer dyspepsion; non-cardiac chest pain;
KW endoscopy; duodenal intubation; ds.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   1..1161
FT   /*tag= a
FT   /product= "MTL-R1B"
XX
PN WO9964436-A1.
XX
PD 16-DEC-1999.
XX
PF 08-JUN-1999; 99WO-US12773.
XX
PR 12-JUN-1998; 98US-0089098.
XX
PA (MERI) MERCK & CO INC.
XX
PI Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
PI Pong S, Smith RG;
XX
DR WPI; 2000-105868/09.
DR P-PSDB; AAY54146.
XX
PT Novel receptor protein for screening compounds used in treating
PT irritable bowel syndrome, constipation and other gastric conditions
XX
PS Claim 6; Fig 4; 44pp; English.
XX
CC The present sequence encodes splice variant MTL-R1B of the motilin
CC receptor. The gene encodes a G-protein coupled receptor, and is
CC designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,
CC MTL-R1A (see AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a
CC functional seven transmembrane domain form, and MTL-R1B is a truncated
CC five transmembrane domain. The MTL-R1 proteins are used to identify
CC agonists and antagonists which can be used for treating gastric motility
CC disorders, functional defects, disorders secondary to neurological
CC disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
CC dysmotility, diabetes, infections, stress-related motility disorders,
CC psychgenic disorders, gastroparesis, gastro-oesophageal reflux disease,
CC constipation, chronic idiopathis pseudo obstruction, acute faecal
CC impaction, postoperative ileus, gallstones, infantile collic, irritable
CC bowel syndrome, non-ulcer dyspepsion, non-cardiac chest pain and
CC diarrhoea. They can also be used in the preparation for colonoscopy,
CC endoscopy and duodenal intubation.
XX
SQ Sequence 1390 BP; 190 A; 496 C; 436 G; 268 T; 0 other;

Query Match   18.5%;   Score 134.8;   DB 21;   Length 1390;
Best Local Similarity   56.1%;   Pred. No. 2.1e-20;
Matches   275;   Conservative   0;   Mismatches   212;   Indels   3;   Gaps   1;

```
Qy  128 tcccgtgtgtcgtgtatgtgcaatttttgtggttggggtgtcattggcaatgtcctgg 187
        | |||| | ||| ||||||  |  || ||| ||||| |||  || |
Db  116 tgccggtgaccgctgtgctggtcttcgtctcggggtgaggcaaacgtggtga 175

Qy  188 tgtgcctggttgattctgcagcaccaagctatgaagacgcccacactacctcttca 247
        ||  |||||  ||| |||  | ||  |||| |||| |||||  ||||| ||||
Db  176 ccgtgaatgctgatcgggcgctaccgggacatgcggaccacaccaacctgtacct 235

Qy  248 gcctggcggtctctgacctcctgttcctggaatgcccctggaggtctatgaga 307
        || ||||| |||| ||| ||| ||||||| || ||||| ||||| ||| ||
```

RESULT 12
ID   AA245402 standard; DNA; 3066 BP.
XX
AC   AA245402;
XX
DT   27-MAR-2000 (first entry)
XX
DE   Genomic sequence of the motilin receptor gene including 5' UTR.
XX
KW   Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
KW   spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
KW   functional defect; neurological disorder; scleroderma; colonoscopy;
KW   paraneoplastic syndrome; radiation induced motility disorder; diabetes;
KW   infection; stress-related motility disorder; psychgenic disorder;
KW   gastroparesis; gastro-oesophageal reflux disease; constipation;
KW   chronic idiopathis pseudo obstruction; infantile collic; diarrhoea;
KW   postoperative ileus; gallstones; infantile collic; diarrhoea;
KW   irritable bowel syndrome; non-ulcer dyspepsion; non-cardiac chest pain;
KW   endoscopy; duodenal intubation; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FH
FT   misc_feature   1929..1931
FT   /*tag= a
FT   /note= "imperfect donor site"
FT   intron   1930..2728
FT   /*tag= b
FT   /note= "intronic sequence"
FT   misc_feature   2080..2082
FT   /*tag= c
FT   /note= "perfect donor site"
FT   misc_feature   2729..2732
FT   /*tag= d
FT   /note= "perfect splice acceptor site"
XX
PN   WO9964436-A1.
XX
PD   16-DEC-1999.
XX
PF   08-JUN-1999; 99WO-US12773.
XX
PR   12-JUN-1998; 98US-0089098.
XX
PA   (MERI) MERCK & CO INC.
XX
PI   Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
PI   Pong S, Smith RG;
XX
DR   WPI; 2000-105868/09.
XX
PT   Novel receptor protein for screening compounds used in treating
PT   irritable bowel syndrome, constipation and other gastric conditions
XX
PS   Example 1; Fig 1; 44pp; English.
XX
CC   The present sequence represents the genomic sequence of the motilin
CC   receptor gene, including the 5' untranslated region (5' UTR). This gene
CC   encodes a G-protein coupled receptor, and is designated MTL-R1 (also
CC   GPR38). Two spliced forms of GPR38 exist, MTL-R1A (see AAY54145) and
CC   MTL-R1B (see AAY54146). MTL-R1A is a functional seven transmembrane
CC   domain form, and MTL-R1B is a truncated five transmembrane domain. The
CC   MTL-R1 proteins are used to identify agonists and antagonists which can
CC   be used for treating gastric motility disorders, functional defects,
CC   disorders secondary to neurological disorders e.g. scleroderma,
CC   paraneoplastic syndromes radiation induced dysmotility, diabetes,
CC   infections, stress-related motility disorders, psychgenic disorders,
CC   gastroparesis, gastro-oesophageal reflux disease, constipation,
CC   chronic idiopathis pseudo obstruction, acute faecal impaction,
CC   postoperative ileus, gallstones, infantile collic, irritable bowel
CC   syndrome, non-ulcer dyspepsion, non-cardiac chest pain and diarrhoea.
CC   They can also be used in the preparation for colonoscopy, endoscopy and
CC   duodenal intubation.
XX
SQ   Sequence 3066 BP; 585 A; 930 C; 926 G; 625 T; 0 other;

PR   29-SEP-1999; 99US-0156634.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Behan DP, Lehmann-Bruinsma K, Chalmers DT, Chen R, Dang HT;
PI   Gore M, Liaw CW, Lin I, Lowitz K, White C;
XX
DR   WPI; 2000-317986/27.
DR   P-PSDB; AAB02854.
XX
PT   Non-endogenous, human G protein-coupled receptors for screening
PT   receptor, inverse or partial agonists useful as therapeutic agents
XX
PS   Example 2; Page 166-168; 187pp; English.
XX
CC   The present invention describes transmembrane receptors, preferably
CC   human G protein coupled receptors (GPCR), for which the endogenous
CC   ligand is unknown (orphan GPCR receptors). More specifically the present
CC   invention relates to non-endogenous, constitutively activated versions
CC   of a human GPCR. These non-endogenous human GPCRs can be useful for
CC   the direct identification of candidate compounds as receptors agonists,
CC   inverse agonists or partial agonists for use as pharmaceutical agents.
CC   AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
CC   the exemplification of the present invention.
XX
SQ   Sequence 2040 BP; 350 A; 630 C; 597 G; 463 T; 0 other;

Query Match   18.5%; Score 134.8; DB 21; Length 2040;
Best Local Similarity   56.1%; Pred. No. 2.3e-20;
Matches 275; Conservative   0; Mismatches 212; Indels   3; Gaps   1;

```
Qy   128 tccccgtgtctgtatgtgcaattttgggtgggggtcaattgtcctcg 187
         ||| ||| ||| ||  ||| ||||||||| | |||||| ||
Db   116 tgtgcgtgaccgctgtgtgccctgttcgtcggtgggtgggcaacgtggtga 175

Qy   188 tgtgctgtggtgattctgcagcacacaggctgtgaagacgcccaacaactactcttca 247
         |  ||| ||  |  |||||  ||||| |||||||| |||  ||  |||
Db   176 ccgtgatgctgatcgggcgctacggagacaatgcggaccacacaacttgtactggca 235

Qy   248 gcctggcggtctctgacctcctgctcctccctggaatgccctgggagtctgtatgaga 307
         ||  |||| ||  ||||||| ||| ||  |||| |||||| ||||
Db   236 gcatggccgtgtccgacctactcatcctggctc---gggctgcgttcgacctgtaccgcc 292

Qy   308 tgtgtggcgcaactaccctttcttgttcgggcccgtgctacttcaagagacggccctct 367
         ||  ||| ||||||| |||  |||   ||||||
Db   293 tctggccgtccgcgcccctgggtgttcggggccgcggcgcctgtccctctacggtgg 352

Qy   368 ttgagaccgtgtcttcgcctccatccatcacacacgtcaagcgtgagcgctacg 427
         ||| || ||||| ||||||||  ||||| |||||||||
Db   353 gcgagggctgcaactacagctgtgtcacatgaccgacgtcgagcgctacc 412

Qy   428 tggccatcctacacaccgttccgcgccaaactgcagagacgacaccgcgccgggcccctcagga 487
         |||||||||  |||| ||  ||||| ||||  |||  ||||
Db   413 tggccatctgccgcgctccgcgccgcgccgcctcggtcaccccggccgccgtccgcgcgc 472

Qy   488 tcctccggcatcgtctgggggcttcctccgtgcttctcctgcccaacaccagcaatccatg 547
         ||  ||| ||  ||  |||  |  |||  ||||| ||  ||
Db   473 tcatcgctgtgtctgggccgtggcgtctctcgccggtctctgttcctgtggtgg 532

Qy   548 gcatcaagttccactacttcccccaatgggtccaggttcggccacctgtacg 607
         |||  ||| ||||| ||  ||||
Db   533 gcgtcgagcgaggaccccggcatctcgtagtcccgggcctcaatgcaccgcggatcg 592

Qy   608 tcatcaagcc 617
         ||  |||
Db   593 cctcctcgcc 602
```

Query Match
Best Local Similarity   18.5%;   Score 134.8;   DB 21;   Length 3066;
56.1%; Conservative   0;   Pred. No. 2.5e-20;   Mismatches 212;   Indels 3;   Gaps 1;
Matches 275; Conservative

Qy   128 tccccgtgtctggtgtatgtgcaattttgttggtggggtcattggccctgg 187
         | || ||||| | | ||| | |||| || |||| |||| | | |||| ||
Db 1144 tgccggtgaccgctgtgtcctggcctgctcgtcggtcggtgagcggcaacgtggtga 1203

Qy   188 tgtgcctggtgtcctgattctgcagcaccaaggctatgaagacagcgcccaacaactactactctca 247
         | ||| || ||| ||| || | | | | ||| | ||||| ||| || ||
Db 1204 ccgtgatgctggatcgggggcgctaccggggaccatgcgggaccacacaacttgtaccctgggca 1263

Qy   248 gcctggcgcggtctctgacctcctggaatgccctgggggttctatgaga 307
         |||||||| | |||| || ||| |||| ||| |||| | ||||
Db 1264 gcatggcgtgtccgacctactcatcctgctc---gggctgcgttcgacctgtaccgcc 1320

Qy   308 tgtggcgcaactaccctttcttgttcggggccgtggggctgctacttcaagacggccctct 367
         ||||||| | | || | ||||| || | | ||| | || | | ||
Db 1321 tctgtggcgttctcgggcgtctgggtgttcggggccgcgctcctggtcccctctcactgg 1380

Qy   368 ttgagacgtgtgttggccgcctccatcctcagcatcacacaccgtcagcgttggagcgctacg 427
         |||| ||| || |||||| | | | ||| || | | ||| || | | ||||
Db 1381 gcgagggctgcaactcaactgctgcacaactgaccggcgctcagcgtcgaagcgctacc 1440

Qy   428 tggccatcctacaccgttccgcgccaaactgcagagcacccggcgcgccggccctcagga 487
         ||||||| | || | || || |||| || | | | | | ||| | ||| ||
Db 1441 tggccatctgccgccgcccgctcctggttgcaccccgccgccgtcccgcggc 1500

Qy   488 tcctcggccatcgtctgggggttcttccggtgtcttctcctgcccaacaccagcatccatg 547
         |||||| |||| ||||| | ||| || || | | |||||| ||| | | ||
Db 1501 tcatcgctgctctgggcgtcgctctctgccggtcccttcttgttcctggtgg 1560

Qy   548 gcatcaagttcccactctcccccaatgggtccccggttcggccactgtacg 607
         |||| | ||| ||| || | | | || || ||| | |||| || |||
Db 1561 gcgtcgagcaggaccccgggatccccggggctccaatggcaccgcgcggatcg 1620

Qy   608 tcatcaagcc 617
         | |||
Db 1621 cctcctcgcc 1630

ID   AA245405 standard; DNA; 1092 BP.
XX
AC   AA245405;
XX
DT   27-MAR-2000 (first entry)
XX
DE   DNA encoding the puffer fish motilin receptor gene clone 75E7.
XX
KW   Motilin receptor; G-protein coupled receptor; MTL-R1; MTL-R1A; MTL-R1B;
KW   spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
KW   functional defect; neurological disorder; scleroderma; colonoscopy;
KW   paraneoplastic syndrome; radiation induced dysmotility; diabetes;
KW   infection; stress-related motility disorder; psychogenic disorder;
KW   gastroparesis; gastro-oesophageal reflux disease; constipation;
KW   chronic idiopathis pseudo obstruction; acute faecal impaction;
KW   postoperative ileus; gallstones; infantile colic; diarrhoea;
KW   irritable bowel syndrome; non-ulcer dyspepsion; non-cardiac chest pain;
KW   endoscopy; duodenal intubation; ds.
XX
OS   Spheroides nephelus.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1092
FT   /*tag= a
FT   /product= "motilin receptor"
XX
PN   WO9964436-A1.
XX
PD   16-DEC-1999.

PF   08-JUN-1999.
XX
PR   12-JUN-1998;   99WO-US12773.
XX
PR   12-JUN-1998;   98US-0089098.
XX
PA   (MERI) MERCK & CO INC.
XX
PI   Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
PI   Pong S, Smith RG;
XX
DR   WPI; 2000-105868/09.
DR   P-PSDB; AAY54147.
XX
PT   Novel receptor protein for screening compounds used in treating
PT   irritable bowel syndrome, constipation and other gastric conditions -
PS   Example 6; Fig 7; 44pp; English.
XX
CC   The present sequence encodes a motilin receptor (clone 75E7) which
CC   is related to the human motilin receptor of the invetion. The
CC   specification describes a G-protein coupled receptor, designated
CC   MTL-R1 (also GPR38). Two spliced forms of GPR38 exist, MTL-R1A (see
CC   AAY54145) and MTL-R1B (see AAY54146), MTL-R1A is a functional seven
CC   transmembrane domain form, and MTL-R1B is a truncated five
CC   transmembrane domain, The MTL-R1 proteins are used to identify agonists
CC   and antagonists which can be used for treating gastric motility
CC   disorders, functional defects, disorders secondary to neurological
CC   disorders e.g. scleroderma, infections, stress-related motility
CC   dysmotility; diabetes, infections, paraneoplastic syndromes radiation
CC   psychgenic disorders, gastroparesis, gastro-oesophageal reflux disease,
CC   constipation, chronic idiopathis pseudo obstruction, acute faecal
CC   impaction, postoperative ileus, gallstones, infantile colic, irritable
CC   bowel syndrome, non-ulcer dyspepsion, non-cardiac chest pain and
CC   diarrhoea. They can also be used in the preparation for colonoscopy,
CC   endoscopy and duodenal intubation.
XX
SQ   Sequence 1092 BP; 205 A; 391 C; 290 G; 206 T; 0 other;

Query Match
Best Local Similarity   18.3%;   Score 133.2;   DB 21;   Length 1092;
51.4%; Conservative   0;   Pred. No. 4.4e-20;   Mismatches 313;   Indels 3;   Gaps 1;
Matches 334; Conservative

Qy   75 caccgaggagtatctggccttcctgcggacctcggcgcagcacttctcagcctggc 134
         |||| | | |||| | || ||| ||||| || ||| || | |||||||
Db   63 caccacggacgacccactactgaggggctccctctccccgcgtccacctcatcccgt 122

Qy   135 gtctgtggttgtatgtgccaattttgttggtggggtcattggccaatgtccggttgcct 194
         || | | | || | | |||||| ||| | | | | ||||| || ||||
Db   123 cacggtcaatctgcatcctcatctggtcggtgtcggtcggacacgtgacctgctggc 182

Qy   195 ggtggttctgcagcaccaaggctatgaagacagcgcccaacaactactactctcagcctggc 254
         || | |||||| | || | || | | ||| | | | | | | ||||
Db   183 catcatccagactcaaggagaacatcaccggccctgaacctgaacgctacctggc 242

Qy   255 ggtctctgacctcctggtctcctggaatgccccctgggaggtctatgaagatggcg 314
         ||||| | | | | | ||| | ||| || | ||| | |||
Db   243 cgtgtccgacctc---gtcatcttcctctgcctcttgctcctgaccctgtacgcctggaa 299

Qy   315 caactaccctttcttgttcggggccgtggggctgctacttcaagacggccccttcttggac 374
         || || ||||| || | | | ||| | || | | | | | || | | ||
Db   300 gtacgtgcggctgtccgcgggcttgccgcctccacactacttcgaagg 359

Qy   375 cgtgtgcttggccgcctccatcctcagcatcacacaccgtcagcgtggacgctacgtggccat 434
         | | | | |||| | | | ||| || | | |||| | | || | || | ||||
Db   360 ctgcacgtcggcgccaccatctccacatcacccggtcaccggatcgaagctacctggccat 419

Qy   435 cctacaccgttccgcgccaaactgcagagacccggcgcgccgggccctcaggatcctcg 494
         ||| || | || || |||| | | | | || || || | | || | || || ||
Db   420 cagcttccccctccaggagcaagttgatggtgaccaggagagggttccagtacacatcatcct 479

Qy   495 catcgtctgggcttcccgtgtctctcctgcccaacaccagcatccatggcatcaa 554
         | |||| || |||| |||| | | | ||| | || || | | || | ||
Db   495 catgtctgggtcttcccgtgcctcttttcttgttcctggtgacaggagaaggtccagtacacaatcct 554

Db 480 ggccctgctggtgtcgctctcgccctgctggtcggcgcctcttctggtcgggtgga 539
Qy 555 gttccactacttccccaatgggtccctggtccccactgtacggtcatcaa 614
Db 540 gtacgacacgagacgacccccgactacaacactgtcaccatcaa 599
Qy 615 gcccatgtggatctacaaatttcatccatccaggtcaacctcctattctacctcctccc 674
Db 600 catcagctcgggggcagctgcacaatcatcggtgtcc 659
Qy 675 catgactgtcatcaagtgtcctctactacctcatggcactcagagtgagta 724
Db 660 gatgctgtcgtctcctcttcacggtgctccatcgggtgcaagtgtgtgga 709

RESULT 14
AAT33127
ID   AAT33127 standard; cDNA; 1254 BP.
XX
AC   AAT33127;
XX
DT   07-NOV-1996   (first entry)
XX
DE   Human neurotensin receptor encoding sequence.
XX
KW   neurotensin receptor; antagonist; agonist; screening; treatment; ulcer;
KW   Parkinson's disease; depression; dementia; retrograde oseophagitis; ds.
XX
OS   Homo sapiens.
XX
PN   JP08143597-A.
XX
PD   04-JUN-1996.
XX
PF   24-NOV-1994;   94JP-0289882.
PR   24-NOV-1994;   94JP-0289882.
XX
PA   (TAKE) TAKEDA CHEM IND LTD.
XX
PI   Claim 4; Page 22; 26pp; Japanese.
XX
PT   DNA encoding human neurotensin receptor protein - useful for
PT   screening for (ant)agonists used to treat, e.g. Parkinson's disease,
PT   depression, dementia, retrograde oseophagitis, ulcers, etc.
XX
DR   WPI; 1996-318958/32.
DR   P-PSDB; AAR98562.
XX
CC   The present sequence encodes a human neurotensin receptor protein
CC   (AAR98562). The DNA and protein are useful for screening for
CC   (ant)agonists used to treat Parkinson's disease, depression, dementia,
CC   retrograde oseophagitis, ulcers, etc..
XX
SQ   Sequence 1254 BP; 215 A; 459 C; 359 G; 221 T; 0 other;

Query Match   17.48%; Score 127; DB 17; Length 1254;
Best Local Similarity 53.98%; Pred. No. 9.9e-19;
Matches   337; Conservative   0; Mismatches 270; Indels 18; Gaps   3;

Qy 114 cagccacttcttcctcccctgtctggtgtatgtgccaattttttgtgggggtcat 173
Db 177 catctactaccaaggttgctgtgttgaccgccgtcttcggctctgtggccacggt 236
Qy 174 tggcaatgtcctggtgtgcctggtagttcgcagcaccaggtctatgaagacgc------ 226
Db 237 gggcaacaggtgacggcgtcacgctgcggaagagaagtcgctgcgagacctgcaagag 296
Qy 227 --ccaccaactactaccttcttcagcctggacctctggtcctcctgg 284
Db 297 cacggtgcattcaccacctgggcagcctggcctgtccgacctgctgctggc 356

Qy 285 aatgccctgctggaggtctatgag---atgtggcgcaaactacccttctttgttcgggccgt 341
Db 357 catgccgctggagctgtacaactccatctgggtgcaccaccctgggccttcggcgcagc 416
Qy 342 gggctgctacttcaagacgggccctcttgagacgcgttgcttcgcctccatccagcat 401
Db 417 cggctgcgcgcggctactacttcctgcgcgagcggctgcacacggcgcgccctcaagt 476
Qy 402 caccaccgtcagcgtgggagcgtacgtggccatcctacaaccgttccgcgcaaactgga 461
Db 477 ggccagctgagtgtggagcgctaacctggccaactgcacccccttcaagccaagaccct 536
Qy 462 gagcaccgggacgcggcgcccccaggatcctcggcgatcctgtctgggcttctcgtgctt 521
Db 537 catgtcccgaagccgcaccaagaagttcatcagcgccatctggctcggccctgcct 596
Qy 522 ctcctgccaacaccagcaatgatcaagttccactacttcccaatgggtccct 581
Db 597 ggcggtgctatgctgttcaccatgggcgagcagagaccgcagcgcgacgt------cca 650

RESULT 15
AA261492
ID   AA261492 standard; cDNA; 1050 BP.
XX
AC   AA261492;
XX
DT   19-JUN-2000   (first entry)
XX
DE   cDNA encoding canine growth hormone secretagogue receptor.
XX
KW   Canine; growth hormone secretagogue receptor; GHSR; dog; ss.
XX
OS   Canis sp.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1050
FT   /*tag= a
FT   /product= "growth hormone secretagogue receptor"
XX
PN   WO20000009538-A2.
XX
PD   24-FEB-2000.
XX
PF   06-AUG-1999;   99WO-US17915.
PR   10-AUG-1998;   98US-0095960.
XX
PA   (MERI) MERCK & CO INC.
XX
PI   Howard AD, Palyha OC, Smith RG, Tan CP;
XX
DR   WPI; 2000-224272/19.
DR   P-PSDB; AAY69293.
XX
PT   A novel canine growth hormone secretagogue receptor used to identify
PT   agonists and antagonists -
XX
PS   Claim 2; Fig 1; 32pp; English.
XX
CC   The present sequence encodes a canine growth hormone secretagogue

Qy 642 ccaggtcacctccttctattctcctccccatgactgtcatcagtgtctctacta 701
Db 651 gcacgccggcggcctggtcgtgcaccccaacatccacaactgcaaggtcgtcat 710
Qy 702 cctcatgcgcactcaagagtgagtatc 726
Db 711 acagtcaacacttcatgtctccatattccccatgtggttgtcatctggtcctgaacac 770
Qy   cctcatgcgcactcaagagtgacgtc 795
Db 771 catcatgcgcaacaagctgaccgtc 795

```
CC   receptor (GHSR). The GHSR cDNA sequence was isolated from a canine
CC   cDNA library constructed from spleen. The GHSR polynucleotide is a
CC   source of probes and primers, which can be used to isolate GHSR
CC   genes from other species. It is also used for recombinant expression
CC   of the polypeptide. The GHSR polypeptide is used in assays to
CC   identify agonists, antagonist and inhibitors. Such identified
CC   compounds can serve as leads for the development of veterinary
CC   pharmaceuticals that can be used to treat dogs having illnesses in
CC   which inappropriate activity of the canine GHSR proteins involved. The
CC   polypeptides may also be used to raise antibodies for diagnostic and
CC   therapeutic purposes.
XX
SQ   Sequence 1050 BP; 146 A; 349 C; 330 G; 224 T; 1 other;
```

Query Match   16.5%;   Score 120.4;   DB 21;   Length 1050;
Best Local Similarity   52.9%;   Pred. No. 2.5e-17;
Matches 306;   Conservative   0;   Mismatches 266;   Indels   6;   Gaps   2;

```
Qy    132 cgtgtcctggtgtatgtgccaattttgtggggtcattgcaatgtcctggtgtg 191
          ||||| |      |  ||| ||  |  |||||  || |||||||||  | |
Db     84 cgtgaacggccaacctgcgtggcgtgttccgtgggcgtcgcgggcaactgctgacggt 143

Qy    192 cctggtgattctgcaccaaggctatgaagacgcccaactactacctcttcagcct 251
          |||||||| | ||  || |||  | | | ||||||||| ||| | ||| |||||||
Db    144 gctggtgtcgcgggcgcttccgcgagactcgccaccacaacctgtacctgcagcct 203

Qy    252 ggcggtgtctctgacctcctgtctcctcttggaatgtccctggaggtctatggagtg 311
          ||| || || ||| |  | ||  || ||| |||     | |    ||| | | | | |
Db    204 ggcctgctcga--cctgctcatctcctgtgcatgccgctcgacctggtgcgcctgtg 260

Qy    312 gcgcaactaccctttcttgttcgggccgtgggctgctacttcaagacgccctcttga 371
          | |      |  ||| ||  |  |||||||| |   ||| ||  | |      |   |
Db    261 gcagtaccggccctggacttcgcgacctgctctcgcaaactcttccagttcgtgagcga 320

Qy    372 gaccgtgtgctttcgcctccatcctcaagcatcaccacccgtcagcgtggagcgtacgtggc 431
          | |  ||| | || ||    |  ||| |  ||  |||| |||||  |||  ||| ||| || |
Db    321 gggctgcacctacgcagtgctcaacatcacgcgctgagcgtcagggctacttcgc 380

Qy    432 catcctacacccgttccgcgcccaaaactgcgagagacaccgcgcgggccctcaggatcct 491
          ||| |       ||  ||  | ||| ||| ||| || || |||  | ||||  |||| ||
Db    381 catctgcttcccgctcccgctcgccaagtgctggtgaccaaggcgcgtgaagctggccct 440

Qy    492 cggcatcgtctgggggcttctccgtgctcttctccctgccaacaccagcatccatggcat 551
          ||||  |||  ||||| ||||| | |  || |  |||  |      | ||  | ||||| |
Db    441 gctggccatctggcgccgtggcctctgcgacgcggccacctcttctgctggtgggcgt 500

Qy    552 caagttccactacttcccaatgggtcccctggttcggccacctgtacgtcat 611
          | ||| ||  ||  | |  |  |  |  |  ||||| |||  ||| ||| |||
Db    501 ggagcacgagaacgcacgacccgggacacccggagtgccgcgccaccgagttcgc 560

Qy    612 caagcccatgtgggatctacaaattcatcatccagtgtcacctccttcctattcttctacctcct 671
          | ||| |||||||  |||  |   |||| | || ||| ||| ||  |   | | | |||| ||||
Db    561 cgtgcgctcggggctgctcacggccatggtgttgggtgtccagcgtcttcttctt---tcct 617

Qy    672 cccccatgactgtcatcaagtgtcctactacctcatgg 709
          |  || |||   | || | |  | ||  | |||| |||| |
Db    618 gcccgtcttctgctcacgtgctctacggcctcatcg 655
```

OM nucleic - nucleic search, using sw model

Run on: August 30, 2001, 01:48:10 ; Search time 69.99 Seconds
                 (without alignments)
                 1971.822 Million cell updates/sec

Title: US-09-684-725-1
Perfect score: 729
Sequence: 1 atggaaaaacttcagaatgc............cactcagagtgagtatctag 729

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
             Maximum Match 100%
             Listing first 45 summaries

Database :   Issued_Patents_NA:*
         1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
         2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
         3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
         4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
         5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
         6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	120.4	16.5	1088	4	US-09-077-675A-6	Sequence 6, Appli
2	120.4	16.5	1122	4	US-09-077-675A-9	Sequence 9, Appli
3	114	15.6	1063	4	US-09-077-675A-1	Sequence 1, Appli
4	112.4	15.4	1029	4	US-09-077-675A-4	Sequence 4, Appli
5	112.4	15.4	1161	1	US-08-086-439C-2	Sequence 2, Appli
6	112.4	15.4	1161	1	US-08-434-877-2	Sequence 2, Appli
7	112.4	15.4	1367	3	US-08-475-742-3	Sequence 3, Appli
8	112.4	15.4	1370	1	US-08-056-051-1	Sequence 1, Appli
9	112.4	15.4	1370	1	US-07-928-611-17	Sequence 17, Appl
10	112.4	15.4	1370	2	US-08-487-811A-17	Sequence 17, Appl
11	112.4	15.4	1370	4	US-09-060-694-17	Sequence 17, Appl
12	112.4	15.4	1370	5	PCT-US93-07370-17	Sequence 17, Appl
13	112.4	15.4	1466	1	US-08-056-051-3	Sequence 3, Appli
14	112.4	15.4	1466	1	US-07-928-611-19	Sequence 19, Appl
15	112.4	15.4	1466	2	US-08-487-811A-19	Sequence 19, Appl
16	112.4	15.4	1466	4	US-09-060-694-19	Sequence 19, Appl
17	112.4	15.4	1466	5	PCT-US93-07370-19	Sequence 19, Appl
18	112.4	15.4	1610	1	US-08-056-051-5	Sequence 5, Appl
19	112.4	15.4	1610	1	US-07-928-611-21	Sequence 21, Appl
20	112.4	15.4	1610	2	US-08-487-811A-21	Sequence 21, Appl
21	112.4	15.4	1610	4	US-09-060-694-21	Sequence 21, Appl
22	112.4	15.4	1610	5	PCT-US93-07370-21	Sequence 21, Appl
23	110	15.1	1529	3	US-08-858-876A-3	Sequence 3, Appli
24	109.2	15.0	1092	4	US-09-077-675A-15	Sequence 15, Appl
25	109.2	15.0	3129	4	US-09-077-675A-14	Sequence 14, Appl
26	97.6	13.4	1575	3	US-08-858-876A-1	Sequence 1, Appli
27	88.8	12.2	1601	1	US-08-722-001-7	Sequence 7, Appli
28	88.8	12.2	1997	1	US-08-722-001-27	Sequence 27, Appl
29	88.8	12.2	2004	1	US-08-722-001-11	Sequence 11, Appl
30	88.6	12.2	1639	1	US-08-334-698-5	Sequence 5, Appl
31	88.6	12.2	1639	1	US-08-228-932-5	Sequence 5, Appl
32	88.6	12.2	1639	1	US-08-468-939-5	Sequence 5, Appl
33	88.6	12.2	1639	2	US-08-406-855A-5	Sequence 5, Appli
34	88.6	12.2	1639	2	US-08-722-190-5	Sequence 5, Appl
35	88.6	12.2	1639	3	US-09-244-354-5	Sequence 5, Appli
36	88.6	12.2	1639	3	US-09-206-899-5	Sequence 5, Appl
37	88.6	12.1	1639	5	PCT-US95-04203-5	Sequence 5, Appl
38	88.4	12.1	836	3	US-09-077-675A-11	Sequence 11, Appli
39	87.6	12.0	1342	3	US-08-832-399-1	Sequence 1, Appli
40	87.6	11.9	1342	4	US-09-372-498-1	Sequence 1, Appl
41	87	11.9	1205	1	US-08-417-103-13	Sequence 13, Appli
42	87	11.9	1634	2	US-07-816-283-1	Sequence 1, Appl
43	87	11.9	1634	1	US-08-417-103-1	Sequence 1, Appl
44	86.8	11.9	921	3	US-08-722-001-17	Sequence 17, Appli
45	86.8	11.9	1567	1	US-08-722-001-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-09-077-675A-6
Sequence 6, Application US/09077675A
Patent No. 6242199

GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Railway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1088 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-09-077-675A-6

Query Match   16.5%;   Score 120.4;   DB 4;   Length 1088;

Best Local Similarity 52.3%; Pred. No. 1.6e-17;
Matches 290; Conservative 0; Mismatches 261; Indels 3; Gaps 1;

```
Qy  132 cgtgtctgtgtatgtgccaatttttgtgggggtcattgcaatgtcctggtgtg 191
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  122 CGTCACAGCAACCTGCGTGGCACTCTTCGTGGTGGGTATCGCTGGCAACCTGCTCACCAT 181

Qy  192 cctggtgattctgcagcaccaggctatgaagacgcccacaacaactactacctcttcagcct 251
        ||||||| |||||||||||||||||||||||
Db  182 GCTGGTGGTGTGCGCCGGCCGTTCCCGAGCGTGCGCCACCACAACCTCTACCTGCTCCAGCAT 241

Qy  252 ggcggtgtctctgacctcctgtgtccttcctggaatgccctgaggttctatgagatgtg 311
        |||||| | | |||||||||||| || ||||||||
Db  242 GGCCTTCTCCGAT---CTGCTCATCTTCCCTCTGCATGCCCCTGGACCTCGTTCGGCCTCTG 298

Qy  312 gcgcaactaccctttcttgttcggggcccgtgggctacttcaagacggccctctttga 371
        |||| ||||||||||||
Db  299 GCAGTACCGGCCCTGGGAACTTCGCGGGACTCCTCTGCAAACTTCTTCCAATTCGTCAGTGA 358

Qy  372 gaccgtgtgtcttcgctccatcctcagctcacgttggagcgctacgtggc 431
        ||||| ||| | ||||
Db  359 GAGCTGCACCTACGGCCACGGTGCTCACCATCACACAGCGTTCGACGCGCTACTTCGC 418

Qy  432 catcctacaccgttccgcgccaaactgcagagagcaacccggcgccgtcaggatcct 491
        |||||||
Db  419 CATCTGCTTCCACTCCGGCGGCCAAGGTGGTGGCTCACCAAGGGCGGGTGAAGCTGGTCAT 478

Qy  492 cggcatcgtctggggcttctccgtgctcttctcccctgcccaacacacagcaccaatggcat 551
        |||||||||||||||
Db  479 CTTCGTCATCTGGGCGCGTGGCCTTCTGCACGCGCGCGGCCCATCTTCGTCGTCAGTCGGGGT 538

Qy  552 caagttccactacttcccaatgggtccctggtcccaggttcggcgcacctgtacggtcat 611
        ||||||||
Db  539 GGAGCACGAGAACGGCACCGACCCTTGGGACACACGAGTGCGCGCCCCACCACGAGTTTGC 598

Qy  612 caagcccatgtggatctacaaattcatcatccaggtcacctcctttccatttctacctcct 671
        |||||||
Db  599 GGTGCGGCGTCTTGGACTGCGCTCACGGTGTGTCATGGTGTGGTGGTCCAGCATCTTCTTCTTCCTTCC 658

Qy  672 cccatgactgtca 685
        | | |||
Db  659 TGTCTTCTGTCTCA 672
```

RESULT 2
US-09-077-675A-9
; Sequence 9, Application US/0907767675A
; Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1122 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-077-675A-9

Query Match   16.5%;   Score 120.4; DB 4; Length 1122;
Best Local Similarity 52.3%;   Pred. No. 1.6e-17;
Matches 290; Conservative 0; Mismatches 261; Indels 3; Gaps 1;

```
Qy  132 cgtgtctgtgtatgtgccaatttttgtggtgggggtcattgcaatgtcctggtgtg 191
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  387 CGTCACAGCGACCTGCGTGGCACTCTTCGTGGTGGGTATCGCTGGCAACCTGCTCACCAT 446

Qy  192 cctggtgattctgcagcaccaggctatgaagagcgcccacaacaactactacctcttcagcct 251
        ||||||| |||||||||||||||
Db  447 GCTGGTGGTGTGCGCCGGCCGTTCCCGAGCGTGCGCCACCACAACCTCTACCTGCTCCAGCAT 506

Qy  252 ggcggtgtctgacctcctggtcctgtctccttggaatgccctgggaggtctatgagatgtg 311
        |||||| || ||||||||||| || ||||||||
Db  507 GGCCTTCTCCGAT---CTGCTCATCTTCCCTCTGCATGCCCGGACCTCGTTCGGCCCTG 563

Qy  312 gcgcaactaccctttcttgttcggggcccgtgggctacttcaagacggccctctttga 371
        |||| ||||||||||||
Db  564 GCAGTACCGGCCCTGGGAACTTCGCGGGACTCCTCTGCAAACTTCTTCCAATTCGTCAGTGA 623

Qy  372 gaccgtgtgtcttcgctccatcctcagctcacgttggagcgctacgtggc 431
        ||||| ||| | ||||
Db  624 GAGCTGCACCTACGGCCACGGTGCTCACCATCACACAGCGTTCGACGCGCTACTTCGC 683

Qy  432 catcctacaccgttccgcgccaaactgcagagagcaacccggcgccgtcaggatcct 491
        |||||||
Db  684 CATCTGCTTCCACTCCGGCGGCCAAGGTGGTGGCTCACCAAGGGCGGGTGAAGCTGGTCAT 743

Qy  492 cggcatcgtctggggcttctccgtgctcttctcccctgcccaacacacagcaccaatggcat 551
        |||||||||||||||
Db  744 CTTCGTCATCTGGGCGCGTGGCCTTCTGCACGCGCGCGGCCCATCTTCGTCGTCAGTCGGGGT 803

Qy  552 caagttccactacttcccaatgggtccctggtcccaggttcggcgcacctgtacggtcat 611
        ||||||||
Db  804 GGAGCACGAGAACGGCACCGACCCTTGGGACACACGAGTGCGCGCCCCACCACGAGTTTGC 863

Qy  612 caagcccatgtggatctacaaattcatcatccaggtcacctcctttccatttctacctcct 671
        |||||||
Db  864 GGTGCGGCGTCTTGGACTGCGCTCACGGTGTGTCATGGTGTGGTGGTCCAGCATCTTCTTCTTCCTTCC 923

Qy  672 cccatgactgtca 685
        | | |||
Db  924 TGTCTTCTGTCTCA 937
```

RESULT 3
US-09-077-675A-1
; Sequence 1, Application US/0907767675A
; Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.

; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1063 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-077-675A-1

```
Query Match                15.6%;   Score 114;  DB 4;   Length 1063;
Best Local Similarity      52.2%;   Pred. No. 3.7e-16;
Matches   302;  Conservative    0;  Mismatches  270;   Indels    6;  Gaps    2;

Qy     132 cgtgtctgtggtgtatgtgccaatttttgtggtgggggtcattggcaatgtcctggtgtg 191
               |||||||||||||||||||||  || ||||||  |||   ||  | | |||
Db      97 CGTCACCGCCACCTGCGTGGCGGCTCTTCGTGGTGGGGTATCGCGGGCCAACCTGCTCACGAT 156

Qy     192 cctggtgtgattctgcagcaccagggctatgaagacgcccaccaacactacctctcagct 251
           | ||||| |||||||||||||  |||| |||||||||||||||| ||| |  | | |||||
Db     157 GCTGGTAGTGTCACGCTTCCGCGAGATGCGCACCACCAACCTCTACCTGTGTCCAGCAT 216

Qy     252 gggggtctctgacctcctgctcctggaatgccctgaggtctatgatgtg 311
             ||||||  ||||| ||  || ||  |  |||||| |  | |
Db     217 GGCCTTCTCCGAACCTAC---TCATCTTCCTCTGCCAACCTCCTTCCCGCCTG 273

Qy     312 gcgcaaactacccttctcttgttcgggcccgtgggcgtgctacttcaaagacggcgccctcttga 371
           ||  | |||||  |||   |||   |    |  ||   |||||  ||  |  |||| |
Db     274 GCAGTACCGGCCCTTGGAACCTTGGCAACCTGCTTGCAAACTCTTCCAGTTCGTTAGCGGA 333

Qy     372 gacgtgttgcttcgcctccatcctcagcatcaccacgtcagcgtgaggcgtacgtggc 431
           ||||||||||   |  |    ||| |||  ||  | | || ||  |  || || | |
Db     334 GAGCTGCACCTAGCCACAGTGCTCACCATCACCGCGCGTGAGCGTGCGACGGCTACTTCGC 393

Qy     432 catcctacacaccgttccgccaaactgagagacccgggcccctcaggatcct 491
           ||||      | ||| |  |||||  |  |||| | | | |  ||  | | ||
Db     394 CATCTGCTTCCCCGCTCCGGCCAAGGTAGTGGTCACCAAGGCGCGGGTAAAGCTGGTCAT 453

Qy     492 cggaatcgtcgtgggggcttctctcctgtcccaacaccagcatcatggcat 551
           ||  || || ||||||| ||||| |   ||||||||||  |||| ||||||
Db     454 CCTGGTCATCTGGGGCCGCGGCCGCTGGTGGGGCGTGGGCCCATCTCTTCGTGCTGGTCGGGAGT 513

Qy     572 ccccatgactgtcatcagtgtcctctactcatgg 709
             || |   | | |  |  | |||||  || |  |  |||
Db     631 GCCTGCTGCTTTCTTCGCCCTCACTGTTGCCTCACTGTTGCCTCTTATAGCCTCATCG 668
```

RESULT   4
US-09-077-675A-4
; Sequence 4, Application US/0907675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1029 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-077-675A-4

```
Query Match                15.4%;   Score 112.4;  DB 4;   Length 1029;
Best Local Similarity      52.1%;   Pred. No. 8.1e-16;
Matches   301;  Conservative    0;  Mismatches  271;   Indels    6;  Gaps    2;

Qy     132 cgtgtctgtggtgtatgtgccaatttttgtggtgggggtcattggcaatgtcctggtgtg 191
               |||||||||||||||||||||  || ||||||  |||   ||  | | |||
Db     192 CGTCACCGCCACCTGCGTGGCGGCTCTTCGTGGTGGGGTATCGCGGGCCAACCTGCTCACGAT 251

Qy     192 cctggtgattctgcagcaccagggctatgaagacgcccaccaacactacctctcagct 251
           | ||||| |||||||||||||  |||| |||||||||||||||| ||| |  | | |||||
Db     294 CGTCACCGCCACCTGCGTGGCGGCTCTTCGTGGTGGGGTATCGCGGGCCAACCTGCTCACGAT 353

Db     354 GCTGGTAGTGTCACGCTTCCGCGAGATGCGCACCACCAACCTCTACCTGTGTCCAGCAT 413
```

RESULT 5
US-08-086-439C-2
Sequence 2, Application US/08086439C
Patent No. 5468615

GENERAL INFORMATION:
APPLICANT: Chio, Christopher L.
APPLICANT: Huff, Rita M.
TITLE OF INVENTION: A Synthetic Gene for D4 Dopamine
TITLE OF INVENTION: Receptors
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Upjohn Company, Corp. Intellectual
ADDRESSEE: Property Law
STREET: 301 Henrietta Street
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/086,439C
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Darnley Jr., James D.
REGISTRATION NUMBER: 33,673
REFERENCE/DOCKET NUMBER: 4700
TELECOMMUNICATION INFORMATION:
TELEPHONE: 616-385-5210
TELEFAX: 616-385-6897
TELEX: 22440l
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1161 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-086-439C-2

Query Match   15.4%; Score 112.4; DB 1; Length 1161;
Best Local Similarity 53.4%; Pred. No. 8.2e-16;
Matches 236; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

```
Qy    139 gtggtgtatgtgcaatttttgttggttggggttcattggcaatgtgtcctggtgtgcctggtg 198
          ||||  ||| ||||  ||  | | |||   |||  ||| |||||||||||||||  ||  || ||
Db    112 GTGGGGGGGGGGTGGCTGCTGCTCATCGGCGCGGCGTGCCTGCGCGGGAACTCGCGTCGTTGCGGTGAGC 171

Qy    199 atctgcagcaccaggctatgaagacgcccacaacaactactacctcttcaagcctgcggtc 258
          |  ||||||| |  ||| |  |||||  |||||   ||| ||  |||  ||| ||||||| |
Db    172 GTGGCCACCGACGGCGCCCTGCAGACCCCACCAAACTCCTTACTCCTCATCGTGAGCGTGGGCGGCC 231

Qy    259 tctgacctcctgtctcgctcctttggaatgccctggagttctatgagatggggcaac 318
          ||| | | |||| ||  || |||   |||| || || | |  || | ||| || |||
Db    232 GCCGACCTCCTCCTCGCTCCTCCTGGTCGTCGTCCTGGTGCTGCTCGTCACTCCGAGGTCCAGGGT 291

Qy    319 taccctttcttgttcggggccgtgggggtctgctgacttcaagagacggccctcttggagaccgtg 378
          ||||   || ||  ||  | |||  | |  ||| ||    |||  | | |||  |  |  ||| | |
Db    292 GGGCGGCGTGGCGCTGTGAGCCCCGTCGTGGACGCCCCGTTACATGGACGTGTCATCGCTG 351

Qy    379 tgcttcgcctccatctcagcatcacacccgtcagcgtgtggaagcgctacgtgggccatccta 438
          |||  ||| |||| || |||    |||| ||| |  || |||||  |||||||||||| ||
Db    352 TGCACCGGCTCCAATCTTCAACCTGTCGGCCATCAGCGTGGACAGGTACGTGGCCGTGGCC 411

Qy    439 caccgttcccgcggcaaaactgcagagcacccggcccggggccctcaggatcctcggccatc 498
          |||  ||| ||| ||     |||  | |||||  |||  || |||| ||| |||| ||||||
Db    412 GTGCCGGCTGCGGCTACCAACCGGCAGAGGGTGGGACCGGCAGCTGCTGCTCATCGGCGC 471

Qy    499 gtctgggggcttctccgtgctctccctgcccaaccaccagcatccaatggcatcaagttc 558
          |||||  |  ||||||||||| ||||||  |  |||||||||   ||||| |||||||||
Db    472 ACGTGGCTGCTGCTCTCCCCGCGCGGGGGGCCCGTACTGTGCGCCATGCCCAACGACGTGGGC 531

Qy    559 cactactttcccaatgggtcc 580
          |||||| ||||| |||||| |
Db    532 GGCCGCGGACCCCGCCGTGTGCC 553
```

RESULT 6
US-08-434-877-2
Sequence 2, Application US/08434877
Patent No. 5721132

GENERAL INFORMATION:
APPLICANT: Chio, Christopher L.
APPLICANT: Huff, Rita M.
TITLE OF INVENTION: A Synthetic Gene for D4 Dopamine
TITLE OF INVENTION: Receptors
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Upjohn Company, Corp. Intellectual
ADDRESSEE: Property Law
STREET: 301 Henrietta Street
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49001
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette (DS, HD)
COMPUTER: Gateway 2000, P5-90
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,877
FILING DATE: 1 July 1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Darnley Jr., James D.
REGISTRATION NUMBER: 33,673

(continued alignment block)

```
Qy    252 ggcggtctctgacctcctggtcctgctccttggaatgccctgggagtctatgaggatgtg 311
          ||||         |||||     |||||||  || | | |    | | || | | | | |
Db    414 GGCCTTCTCCGAACTAC---TCATCTTCCTCTGCAATGCCCCTGCGAACTCTTCGCCCTTG 470

Qy    312 gcgcaactaccctttcttgttcggggccgtggcgtgctacttcaagacgacggccctcttga 371
          | | |  |   ||| | ||| |||   |  || |  |  |||   | | | |  ||| |
Db    471 GCAGTACCGGCCGTTGGAACCTTGGACCTTGCGCTCTGCAAACTCTTCCAGTTCGGTTAGCGGA 530

Qy    372 gaccgtgtgcttcgcctccatcccacaccgtcaagccgcctacgtggc 431
          ||  | ||||| | |  |  |  | |||  |  || | |  | | |  |
Db    531 GAGCTGCACCTACCGCACCAGTGTCTCACATACCTCGACGCGCTACTTCGC 590

Qy    432 catcctacaccgttccgcgccaaactgcgagagcaccgcggcgcgggcctcaggatcct 491
          |||  |||  |  ||| ||| |||  | || | || || |  |  ||  ||||| |||
Db    591 CATCTCGCTTCCCGCTGCGGGGCCAAGGTAGTAGGTCACCAAGGGCCGGGTAAAGCTGGTCAT 650

Qy    492 cgggaatcgtctggggcttctccgtgctcttctccctgcccaaccagcatccatggcat 551
          |||  || | | || |||| |||| | || ||   || ||   |||| ||| | | |
Db    651 CCTGGTCATCTGGGCGGTGGCCGTTCTGACGCCATCTTCTGCAGACGCCTTCGTTCGGTGGGAGT 710

Qy    552 caagttccactacttccccaatgggtcccctgtgtccaagtgttcggcgccacctgtacggtcat 611
          |||| | |    |||   ||  | |||   | |  ||||   | || | |  |||  | | ||
Db    711 GGAGCATGATAACGGCACTGACCCTCGGGGACACCAGTGCCGCGCGCCGCCCACGGAGTTCGC 770

Qy    612 caagcccatgtggatctacaacattcatcatccaggtcacctccttctattctaccctcct 671
          |||  | || ||  || |  |  |  ||||| | |||| | ||   | ||| ||   ||
Db    771 CGTGCGGCTCCGGGTGCTTACCGGTCGGGTGTGTCCAGTGGTCGTTCTTCTCT---TCCT 827

Qy    672 cccatgactgtcatcagtgtcctcactacttcatgg 709
          ||  | | | | ||   | ||  |   ||| ||||
Db    828 GCCTGTCTCTTCCTCACTGTGCTCTATAGCCTCATCG 865
```

```
;  REFERENCE/DOCKET NUMBER: 4700 DV1
;  TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 616-385-5210
;     TELEFAX: 616-385-6897
;     TELEX: 224401
;  INFORMATION FOR SEQ ID NO: 2:
;  SEQUENCE CHARACTERISTICS:
;     LENGTH: 1161 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     MOLECULE TYPE: cDNA
US-08-434-877-2
```

```
Query Match      15.4%;   Score 112.4;  DB 1;  Length 1161;
Best Local Similarity   53.4%;   Pred. No. 8.2e-16;
Matches  236;  Conservative   0;  Mismatches  206;   Indels   0;  Gaps   0;

y   139 gtggtgtatgtgccaatttttgtggttgggggtcattggggcaatgtcctggttgtgctggttg 198
        ||||| |         |||| |||||||   || || ||||||||||||||
Db  112 GTGGGGGGGTGGTGCTGCTCATCGGCGGCGGGTTGCTCCGGGGAAACTCGCCTCGTTGTGGGTGAGC 171

Qy  199 atttctgcagcaccaggctatgaagagacgcccaccaactacttcatgctggcggtc 258
        |||||||||||||||||| |||| ||    || ||| ||||| || ||||||||||| |
Db  172 GTGGGCCACCGAGCGGCGGGCCCCTTGCAGAACGCCCACCAAACTCCTTCATCGTGAGCCCTGGCGGCC 231

Qy  259 tctgacctcctggtcctgctcctgggaggtgtctatgagatgtggcgcaac 318
        | |||||||         |||  ||   || || |||| || ||||
Db  232 GCCGACCTCCTCCTCGTCCTTCGTCGCTTCGTCTACTCCGAGGTCCAGGGT 291

Qy  319 tacccttttcttgttcgggccgtgggcgtgctactcaagacagcgccctcttgagaccgtg 378
        || ||||||||||||||     |||  |||||||||   ||| || ||||
Db  292 GGCCGCCGTGCCTGCTGCGAGCCCCCGTGTCGCGACCGCCCCCATGGCCATGCATGCTG 351

Qy  379 tgcttcgcctccatcctcagcatcaccacggtcagcgtggagcgctacgtggccatccta 438
        ||| ||||| |||| |||||| || ||   |||||||||| || | |||||| |||
Db  352 TGCACCGGCCTCCATCTTCAACCTGTCGCGCATCACGGTGGACAGGTTCGTTGGCCGTGGCC 411

Qy  439 caccgttccgcgccaaatgcagaagacaccggggccctcaaggatcctcgggcatc 498
        || |||||| ||||||| ||  |     || ||| || | |||||| ||||| |
Db  412 GTGCGCCGCCTACAACCGGCAGGGTGGGAGCCGCTCGCTCATCCGGCGCC 471

Qy  499 gtctggggcttctccgtgctctccctgcccaacaccagcatccatggcacaagttc 558
        | |||||||||| |      ||||  ||   ||||| || ||||
Db  472 ACGTGGCGCTCGTCGTTCCGCGCGCGGGTGGGCGGCCCCGCTTACGTGTCGCGACGACGTGCGC 531

Qy  559 cactacttcccaatgggtcc 580
        | ||||| ||| ||||
Db  532 GGCCGCGACCCCGCCGTGTCC 553
```

```
RESULT  7
US-08-475-742-3
; Sequence 3, Application US/08475742
; Patent No. 6121015
; GENERAL INFORMATION:
;   APPLICANT: O'Malley, Karen L
;   APPLICANT: Todd, Richard D
;   TITLE OF INVENTION: Gene Encoding the Rat Dopamine D4 Receptor
;   FILE REFERENCE: WU 102 CON DIV
;   CURRENT APPLICATION NUMBER: US/08/475,742
;   CURRENT FILING DATE: 1995-06-07
;   EARLIER APPLICATION NUMBER: US 08/261,293
;   EARLIER FILING DATE: 1994-06-16
;   EARLIER APPLICATION NUMBER: US 08/014,013
;   EARLIER FILING DATE: 1993-01-28
;   NUMBER OF SEQ ID NOS: 16
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1367
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
;  FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (1)..(1367)
;  OTHER INFORMATION: D4 Dopamine Receptor cDNA
;  PUBLICATION INFORMATION:
;     AUTHORS: Van tol, H. H.
;     AUTHORS: Bunzow, J. R.
;     TITLE: Cloning of the gene for a human dopamine D4 receptor
;     TITLE: with high affinity for the antipsychotic clozapine
;     JOURNAL: Nature
;     VOLUME: 350
;     PAGES: 610-614
;     DATE: 1991
US-08-475-742-3
```

```
Query Match      15.4%;   Score 112.4;  DB 3;  Length 1367;
Best Local Similarity   53.4%;   Pred. No. 8.4e-16;
Matches  236;  Conservative   0;  Mismatches  206;   Indels   0;  Gaps   0;

Qy  139 gtggtgtatgtgccaatttttgtggttgggggtcattggggcaatgtcctggttgtgctggttg 198
        ||||| |         |||| |||||||   || || ||||||||||||||
Db  218 gtggggggcgtggtgctgctcatcggcggcgggttgctccgggggaactcgcctcgttgtgggtgagc 277

Qy  199 atttctgcagcaccaggctatgaagagacgcccaccaactacttcatgctggcggtc 258
        |||||||||||||||||| |||| ||    || ||| ||||| || ||||||||||| |
Db  278 gtggccaccgagcgcggcccctgcagaacgcccaccaaactccttcatcgtgagccctggcggcc 337

Qy  259 tctgacctcctggtcctgctcctgggaggtgtctatgagatgtggcgcaac 318
        | |||||||         |||  ||   || || |||| || ||||
Db  338 gccgacctcctcctcgtccttcgtcgcttcgtctactccgaggtccagggt 397

Qy  319 tacccttttcttgttcgggccgtgggcgtgctactcaagacagcgccctcttgagaccgtg 378
        || ||||||||||||||     |||  |||||||||   ||| || ||||
Db  398 ggccgccgtgcctgctgcgagcccccgtgtcgcgaccgcccccatggccatgcatgctg 457

Qy  379 tgcttcgcctccatcctcagcatcaccacggtcagcgtggagcgctacgtggccatccta 438
        ||| ||||| |||| |||||| || ||   |||||||||| || | |||||| |||
Db  458 tgcaccggcctccatcttcaacctgtcgcgcatcacggtggacaggttcgttggccgtggcc 517

Qy  439 caccgttccgcgccaaatgcagaagacaccggggccctcaaggatcctcgggcatc 498
        || |||||| ||||||| ||  |     || ||| || | |||||| ||||| |
Db  518 gtgcgccgcctacaaccggcagggtgggagccgctcgctcatccggcgcc 577

Qy  499 gtctggggcttctccgtgctctccctgcccaacaccagcatccatggcacaagttc 558
        | |||||||||| |      ||||  ||   ||||| || ||||
Db  578 acgtggcgctcgtcgttccgcgcgcgggtgggcggccccgcttacgtgtcgcgacgacgtgcgc 637

Qy  559 cactacttcccaatgggtcc 580
        | ||||| ||| ||||
Db  638 ggccgcgaccccgccgtgtcc 659
```

```
RESULT  8
US-08-056-051-1
; Sequence 1, Application US/08056051
; Patent No. 5516683
; GENERAL INFORMATION:
;   APPLICANT: Grandy, David K
;   APPLICANT: Bunzow, James R
;   APPLICANT: Civelli, Olivier
;   APPLICANT: Van Tol, Hubert H.-M.
;   TITLE OF INVENTION: A No. 5516683e1 Human Dopamine Receptor and Uses
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Allegretti & Witcoff, Ltd.
;     STREET: 10 South Wacker Drive, Suite 3000
;     CITY: Chicago
;     STATE: Illinois
;     COUNTRY: USA
;     ZIP: 60606
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
```

Left column

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/056,051
FILING DATE: 19930429
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5516683nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
US-08-056-051-1
FEATURE:
    NAME/KEY: CDS
    LOCATION: 104..1267
MOLECULE TYPE: cDNA
FEATURE:
    NAME/KEY: 3'UTR
    LOCATION: 1268..1370
    STRANDEDNESS: single
    TYPE: NUCLEIC ACID
    LENGTH: 1370 base pairs
SEQUENCE CHARACTERISTICS:
INFORMATION FOR SEQ ID NO: 1:
FEATURE:
    NAME/KEY: 5'UTR
    LOCATION: 1..103
FEATURE:
```

```
Query Match       15.4%;   Score 112.4; DB 1; Length 1370;
Best Local Similarity 53.4%; Pred. No. 8.4e-16;
Matches 236; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

Qy  139 gtggtgtatgtgccaatttttgtggggtcattggcaatgtcctggttgtgctggtg 198
Db  218 GTGGGGGGGGGCGTTGCTGCTCATCGCGCGGGGGGCGTGCTCGGGAGC 277
Qy  199 atttgtcagcaccaaggctatgaagacgcccacacaactactactcttcagcctggcgggtc 258
Db  278 GTGGCCACCGAGCGCGCCCTGCAGAGACCGCCCACCAACTCCTTCATCGTGACGTGCGGGGCC 337
Qy  259 tctgacctcctggtctcgctccttggaatgccctggaggtctatagagatggcgcgcaac 318
Db  338 GCCGACCTCCTCCTGCGCTCGTTCGTTACTCCGCGAGGTCCAGGGT 397
Qy  319 tacccttcttgttcgggccgtgggctacttcaagagacggcccctcttttgagaccgtg 378
Db  398 GGCGGCGTGGGCGTCGTGAGCCGCCCCGCCCCTCATGGACGCCATGGCCATCGCG 457
Qy  379 tgcttcgcctcccatcctcagcgatcacaccgtcagcgtggagcgctacacgtggccatccta 438
Db  458 TGCACCGCCCCTCCATCTTCAACCTGTGTCGCGCCATCAGCGTGGACAGTGTTCGTGGCCGTGGCC 517
Qy  439 caccgttccgcgccaaactgcagagacaccgggcccggggcggccctcaaggatcctcggcatc 498
Db  518 GTGCCGCTGCGCTACAACCGGCAGGGGTGGGAGCCGCGGCGGACGCCGCGGCCTGCTCATCGGCGC 577
Qy  499 gtctggggcttctccgtgctcttctcccgtgcccaacaacagcatccaatggccatcaagttc 558
Db  578 ACGTGCCTGCTGCTGTCTCCGCGGCGGGCGGGGTGGGCGGCGGCCCGTACTGTGCGGCCTCAACGACGTGCGC 637
Qy  559 cactacttcccaatgggtccc 580
Db  638 GGCCGCGGAACCCCGGCCCGTGTGCC 659

RESULT  9
US-07-928-611-17
```

Right column

```
Sequence 17, Application US/07928611
Patent No. 5569601
GENERAL INFORMATION:
APPLICANT: Van Tol, Hubert H.M.
APPLICANT: Civelli, Olivier
TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
    ADDRESSEE: Allegretti & Witcoff, Ltd.
    STREET: 10 South Wacker Drive, Suite 3000
    CITY: Chicago
    STATE: Illinois
    COUNTRY: USA
    ZIP: 60606
COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/07/928,611
    FILING DATE: 19920810
    CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
    NAME: No. 5569601nan, Kevin E
    REGISTRATION NUMBER: 35,303
    REFERENCE/DOCKET NUMBER: 90,1092-B
TELECOMMUNICATION INFORMATION:
    TELEPHONE: 312-715-1000
    TELEFAX: 312-715-1234
    TELEX: 810-221-8317
US-07-928-611-17
FEATURE:
    NAME/KEY: CDS
    LOCATION: 104..1267
MOLECULE TYPE: cDNA
FEATURE:
    NAME/KEY: 3'UTR
    LOCATION: 1268..1370
    STRANDEDNESS: single
    TYPE: NUCLEIC ACID
    LENGTH: 1370 base pairs
SEQUENCE CHARACTERISTICS:
INFORMATION FOR SEQ ID NO: 17:
FEATURE:
    NAME/KEY: 5'UTR
    LOCATION: 1..103
FEATURE:
MOLECULE TYPE: cDNA
```

```
Query Match       15.4%;   Score 112.4; DB 1; Length 1370;
Best Local Similarity 53.4%; Pred. No. 8.4e-16;
Matches 236; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

Qy  139 gtggtgtatgtgccaatttttgtggggtcattggcaatgtcctggttgtgctggtg 198
Db  218 GTGGGGGGGGGCGTTGCTGCTCATCGCGCGGGGGGCGTGCTCGGGAGC 277
Qy  199 atttgtcagcaccaaggctatgaagacgcccacacaactactactcttcagcctggcgggtc 258
Db  278 GTGGCCACCGAGCGCGCCCTGCAGAGACCGCCCACCAACTCCTTCATCGTGACGTGCGGGGCC 337
Qy  259 tctgacctcctggtctcgctccttggaatgccctggaggtctatagagatggcgcgcaac 318
Db  338 GCCGACCTCCTCCTGCGCTCGTTCGTTACTCCGCGAGGTCCAGGGT 397
Qy  319 tacccttcttgttcgggccgtgggctacttcaagagacggcccctcttttgagaccgtg 378
Db  398 GGCGGCGTGGGCGTCGTGAGCCGCCCCGCCCCTCATGGACGCCATGGCCATCGCG 457
Qy  379 tgcttcgcctcccatcctcagcgatcacaccgtcagcgtggagcgctacacgtggccatccta 438
Db  458 TGCACCGCCCCTCCATCTTCAACCTGTGTCGCGCCATCAGCGTGGACAGTGTTCGTGGCCGTGGCC 517
```

RESULT 10
US-08-487-811A-17
Sequence 17, Application US/0847811A
Patent No. 5883226

GENERAL INFORMATION:
APPLICANT: Civelli, Olivier
APPLICANT: Van Tol, Hubert H.M.
TITLE OF INVENTION: A No. 5883226e1 Human Dopamine Receptor and Uses
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,811A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: No. 5883226nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092-L
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX:
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1370 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..103
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 1268..1370
FEATURE:
NAME/KEY: CDS
LOCATION: 104..1267

Query Match   15.4%;   Score 112.4; DB 2;   Length 1370;
Best Local Similarity 53.4%;   Pred. No. 8.4e-16;
Matches 236;   Conservative 0; Mismatches 206; Indels 0; Gaps 0;

Qy   139 gtggtgtatgtgcaatttttgtggtgggggtcattggcaatgtcctggtcgtgctggtg 198
         ||| ||||||| ||| ||| | ||| |||| || |||||||| | | | ||||| |
Db   218 gtggggggcggctgctcttcatcggtcacggccggttcctcgtgtgcgtgctggagc 277

Qy   199 ...
Db   218 ...

Qy   259 ...
Db   278 ...

Qy   319 ...
Db   338 ...

Qy   379 ...
Db   398 ...

Qy   439 caccgttccgccgcaaactgcaagagcaaccccggcgccgccgggccctcaggatccctcggcatc 498
         |||| ||| ||||||||| ||| |||||| |||||||| || || ||||||||| |||
Db   518 GTGCCGCTGCGCTACAACCCGCAGGGTGGGAGCGCCGCAGGTGCTCCTGCTCATCGGCGCC 577

Qy   499 gtctgggggcttctccgtgtctcttccctgcccaacacagcatccatggcatcaagttc 558
         ||||||||| | ||||| ||| |||||||||| ||||||| |||||||| | ||||||||
Db   578 ACGTGGCGTCGTGTCCGCGGGGGTGGGCCCCCGTACTGTGGCCGCCTCAACGACGTGCGC 637

Qy   559 cactacttcccaatgggtccc 580
         |||||||||| |||||| |||
Db   638 GGCGCGGCGACCCGCCCTGCC 659

RESULT 11
US-09-060-694-17
Sequence 17, Application US/09060694
Patent No. 6203998

GENERAL INFORMATION:
APPLICANT: Civelli, Olivier
APPLICANT: Van Tol, Hubert H.M.
TITLE OF INVENTION: A No. 6203998e1 Human Dopamine Receptor and Uses
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,694
FILING DATE: 15-APR-1998
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: No. 6203998nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092-MM
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX:
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1370 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..103

Qy   199 atttctgaagcaccaggctatgaagacgcccaccaacaactactacctcttcagcctggccggtc 258
         ||| |||||||||||| ||||| |||| |||| |||||||| ||||| || |||||||||||||
Db   278 GTGGCCACCGACGCGCCCCTGCCAGAGACCCCCACCAAACTCCTTCATCGTTGAGCGTGGGCGC 337

Qy   259 tctgacctcctggtcctgctcctcttggaatgcccctggagtctatgagatgtggcgcaac 318
         ||| ||| ||||||| ||| ||||| || |||||||||||||| || | |||| |||| |
Db   338 GCCGACCTTCTCTCGCTCTCCTTCGGTCCTGGTGTTCGTCTGCTTCGTCTACTCGGAGTCCAGGGT 397

Qy   319 taccctttctgttcgggccgtgggcgtcgtacttcaagagacgggcccctcttggaaccgtg 378
         | ||| | || | ||| |||||||||| ||||||||||| ||| | | ||||| || ||||
Db   398 GGCGCGTGGCTGCTGACGCCCGCCGCGGGTGGCGGTGGGCGGGCGGCCATGGACGTCATGCTG 457

Qy   379 tgcttcgcccatcctcagcgatcaccaccgtcagcgtggagcgctacgtggccatcta 438
         ||| |||| ||||||| ||| ||||||| |||||||||| |||| |||| ||||||||
Db   458 TGCACCGCCCCTCCATCTTCAACCTTGTCGGCCATCACGGTGGACAGGTTCGTGGCGTGGGC 517

Qy   439 caccgttccgccgcaaactgcaagagcaaccccggcgccgccggggccctcaggatccctcggcatc 498
         |||| ||| ||||||||| ||| |||||| |||||||| || || ||||||||| |||
Db   518 GTGCCGCTGCGCTACAACCCGCAGGGTGGGAGCGCCGCAGGTGCTCCTGCTCATCGGCGCC 577

Qy   499 gtctgggggcttctccgtgtctcttccctgcccaacacagcatccatggcatcaagttc 558
         ||||||||| | ||||| ||| |||||||||| ||||||| |||||||| | ||||||||
Db   578 ACGTGGCGTCGTGTCCGCGGGGGTGGGCCCCCGTACTGTGGCCGCCTCAACGACGTGCGC 637

Qy   559 cactacttcccaatgggtccc 580
         |||||||||| |||||| |||
Db   638 GGCGCGGCGACCCGCCCTGCC 659

```
;     FEATURE:
;       NAME/KEY: 3'UTR
;       LOCATION: 1268..1370
;     FEATURE:
;       NAME/KEY: CDS
;       LOCATION: 104..1267
US-09-060-694-17

Query Match           15.4%;  Score 112.4;  DB 4;  Length 1370;
Best Local Similarity 53.4%;  Pred. No. 8.4e-16;
Matches  236;  Conservative   0;  Mismatches  206;  Indels   0;  Gaps   0;

Qy    139 gtggtgtatgtgccaatttttgtggtgggggtcattggcaaatgtcctggttgtgcctggtg 198
          ||||| ||| ||||  |||| ||||||| ||||||||  ||| |||||||||| |||||||||
Db    218 GTGGGGGGGCGGGTGCCTTGCTCATCGGCGGGGCGGGGAACTCGCTCGTGTGCGGTGAGC 277

Qy    199 atttctgcagcacccaggctatgaagacgcccacacaaactactacctcttcaagcctggcggtc 258
          |||| ||| ||||| ||||||||| || ||| ||||| ||||||||| |||||||| |||| ||||
Db    278 GTGGCCACCGAGCGCGCCCTGCAGACGCCCACCAAACTCCTTCATCGTGACCTGGCCGCC 337

Qy    259 tctgacctcctggttcctgctcctcttgggatgcccctggaggtctatgagagtggcgcaac 318
          ||||| ||||||| ||||||||||||| |||||| |||||||||||  ||||| || |||||
Db    338 GCCGACCTCCTCTCGCGCTCTCCGGTGCTGCCGGGTTCPACTCCGAGGTCCAGGGT 397

Qy    319 tacccttcttgttcgggccgtgggctgctacttcaagacggcccctcttgagacgtg 378
          | ||| |||||| || |||||| |||| || ||||| ||| |||  ||||| |||||||
Db    398 GGCGGCGTGGCTGGTGGAGCCCCGCCCTCATGGCCATGGACGTCATGCTG 457

Qy    379 tgcttcgctccatcctcagcatcacacaccgtcaagcgtggagcggctacgtggccatcca 438
          |||||||||||| |||| |||||||| |||||||||  ||||||  ||| || ||||||
Db    458 TGCACCGGCCTTCCATCTTCAAACCTGTGCGCCATCAGCGTGGACAGGTTCGGTGCCGTGCC 517

Qy    439 caccgttccgcgcccaaactgcagagagcaaccgcgccgggccctcaggatcctcggcatc 498
          | |||||||||  ||||  ||||||||| ||| || ||| ||||||| ||| |||| |||
Db    518 GTGGCGCGTGCGGTACAACCGGCAGGGTGGGAGCGCGCGGGCCCGCGCAGGTGCGTCTCANCGGCGCC 577

Qy    499 gtctgggggcttctccgtgctcttctccctgcccaacacagcatccatggcatcaagttc 558
          ||| ||| || |||| ||||| || |||||| |||||||||||||||||||||||| |||
Db    578 ACGTGGCTGCTGTCCGGCGTGCCGGGGGCGCGGCCCGCCGGGGCGCCCGCGCGGCGCGTGGGCGCGC 637

Qy    559 cactacttcccaaatgggtccc 580
          |||| ||| |||||||||||||
Db    638 GGCCGCGACCCGCCGCGTGTGCC 659
```

```
;     FEATURE:
;       NAME/KEY: CDS
;       LOCATION: 104..1267
PCT-US93-07370-17

Query Match           15.4%;  Score 112.4;  DB 5;  Length 1370;
Best Local Similarity 53.4%;  Pred. No. 8.4e-16;
Matches  236;  Conservative   0;  Mismatches  206;  Indels   0;  Gaps   0;

Qy    139 gtggtgtatgtgccaatttttgtggtgggggtcattggcaaatgtcctggttgtgcctggtg 198
          ||||| ||| ||||  |||| ||||||| ||||||||  ||| |||||||||| |||||||||
Db    218 GTGGGGGGGCGGGTGCCTTGCTCATCGGCGGGGCGGGGAACTCGCTCGTGTGCGGTGAGC 277

Qy    199 atttctgcagcacccaggctatgaagacgcccacacaaactactacctcttcaagcctggcggtc 258
          |||| ||| ||||| ||||||||| || ||| ||||| ||||||||| |||||||| |||| ||||
Db    278 GTGGCCACCGAGCGCGCCCTGCAGACGCCCACCAAACTCCTTCATCGTGACCTGGCCGCC 337

Qy    259 tctgacctcctggttcctgctcctcttgggatgcccctggaggtctatgagagtggcgcaac 318
          ||||| ||||||| ||||||||||||| |||||| |||||||||||  ||||| || |||||
Db    338 GCCGACCTCCTCTCGCGCTCTCCGGTGCTGCCGGGTTCPACTCCGAGGTCCAGGGT 397

Qy    319 tacccttcttgttcgggccgtgggctgctacttcaagacggcccctcttgagacgtg 378
          | ||| |||||| || |||||| |||| || ||||| ||| |||  ||||| |||||||
Db    398 GGCGGCGTGGCTGGTGGAGCCCCGCCCTCATGGCCATGGACGTCATGCTG 457

Qy    379 tgcttcgctccatcctcagcatcacacaccgtcaagcgtggagcggctacgtggccatcca 438
          |||||||||||| |||| |||||||| |||||||||  ||||||  ||| || ||||||
Db    458 TGCACCGGCCTTCCATCTTCAAACCTGTGCGCCATCAGCGTGGACAGGTTCGGTGCCGTGCC 517

Qy    439 caccgttccgcgcccaaactgcagagagcaaccgcgccgggccctcaggatcctcggcatc 498
          | |||||||||  ||||  ||||||||| ||| || ||| ||||||| ||| |||| |||
Db    518 GTGGCGCGTGCGGTACAACCGGCAGGGTGGGAGCGCGCGGGCCCGCGCAGGTGCGTCTCANCGGCGCC 577

Qy    499 gtctgggggcttctccgtgctcttctccctgcccaacacagcatccatggcatcaagttc 558
          ||| ||| || |||| ||||| || |||||| |||||||||||||||||||||||| |||
Db    578 ACGTGGCTGCTGTCCGGCGTGCCGGGGGCGCGGCCCGCCGGGGCGCCCGCGCGGCGCGTGGGCGCGC 637

Qy    559 cactacttcccaaatgggtccc 580
          |||| ||| |||||||||||||
Db    638 GGCCGCGACCCGCCGCGTGTGCC 659
```

```
RESULT 12
PCT-US93-07370-17
Sequence 17, Application PC/TUS9307370
GENERAL INFORMATION:
  APPLICANT:
  TITLE OF INVENTION:  A Novel Human Dopamine Receptor and Uses
  NUMBER OF SEQUENCES:  22
  COMPUTER READABLE FORM:
    MEDIUM TYPE:  Floppy disk
    COMPUTER:  IBM PC compatible
    OPERATING SYSTEM:  PC-DOS/MS-DOS
    SOFTWARE:  Patentin Release #1.0, Version #1.25 (EPO)
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER:  PCT/US93/07370
INFORMATION FOR SEQ ID NO: 17:
  SEQUENCE CHARACTERISTICS:
    LENGTH:  1370 base pairs
    TYPE:  nucleic acid
    STRANDEDNESS:  single
    TOPOLOGY:  linear
  MOLECULE TYPE:  cDNA
  FEATURE:
    NAME/KEY:  5'UTR
    LOCATION:  1..103
  FEATURE:
    NAME/KEY:  3'UTR
    LOCATION:  1268..1370
```

```
RESULT 13
US-08-056-051-3
Sequence 3, Application US/08056051
Patent No. 5516683
GENERAL INFORMATION:
  APPLICANT:  Grandy, David K
  APPLICANT:  Bunzow, James R
  APPLICANT:  Civelli, Olivier
  APPLICANT:  Van Tol, Hubert H.-M.
  TITLE OF INVENTION:  A No. 5516683el Human Dopamine Receptor and Uses
  NUMBER OF SEQUENCES:  6
  CORRESPONDENCE ADDRESS:
    ADDRESSEE:  Allegretti & Witcoff, Ltd.
    STREET:  10 South Wacker Drive, Suite 3000
    CITY:  Chicago
    STATE:  Illinois
    COUNTRY:  USA
    ZIP:  60606
  COMPUTER READABLE FORM:
    MEDIUM TYPE:  Floppy disk
    COMPUTER:  IBM PC compatible
    OPERATING SYSTEM:  PC-DOS/MS-DOS
    SOFTWARE:  Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER:  US/08/056,051
    FILING DATE:  19930429
    CLASSIFICATION:  435
  ATTORNEY/AGENT INFORMATION:
    NAME:  No. 5516683nan, Kevin E
    REGISTRATION NUMBER:  35,303
```

REFERENCE/DOCKET NUMBER: 90,1092-C
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 312-715-1000
  TELEFAX: 312-715-1234
  TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 3:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1466 base pairs
    TYPE: NUCLEIC ACID
    STRANDEDNESS: single
    TOPOLOGY: linear
  MOLECULE TYPE: cDNA
  FEATURE:
    NAME/KEY: 5'UTR
    LOCATION: 1..103
  FEATURE:
    NAME/KEY: 3'UTR
    LOCATION: 1364..1466
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 104..1363
US-08-056-051-3

Query Match   15.4%;   Score 112.4;   DB 1;   Length 1466;
Best Local Similarity   53.4%;   Pred. No. 8.5e-16;
Matches 236;   Conservative 0;   Mismatches 206;   Indels 0;   Gaps 0;

```
Qy  139 gtggtgtatgtgccaatttttgtgtggggtcattggcaatgtcctggttgtgcctggtg 198
            ||||                  |  |   |||                ||| |||||
Db  218 GTGGGGGGGCGTGCTGTCTCATCGGCCCGGTGCTCGCGGGGGAACTCGCTCGTGCGGTGAGC 277

Qy  199 atctgcagcacagggctatgaagagacgcccacacaactactactcttcagcctggcgtc 258
            |  |   |   | |   |  ||  |   |  | |||||   ||||||||||||||||
Db  278 GTGGCCACCGAGCGCGCCCCTGCAGACGCCCACCAACTCCTTCATCGTGACGCTGGGCGGC 337

Qy  259 tctgacctcctggtctctgctcctgggaatgccctgtctatgagagtgtggcgcaac 318
            |  |||||||| |  |      |  |||||||  |||||||||||||  ||||||
Db  338 GCCGACCTCCTCCTCCTGGTGCTCGCGCGTCTTCGTCTACTCCGAGGTCCAGGT 397

Qy  319 taccctttcttgttcgggccgtgggctgctacttcaagacggccctctctttgagaccgtg 378
            |||   |    |  ||| |||  |  |||  |        | ||   | ||||||||||||
Db  398 GGGCGGCGTGGCTGCTGAGCCCCGGCGCCCTGCCCGCCCATGGCCATGGACGTCATGCTG 457

Qy  379 tgcttccgcctccatccagcatcaccacacgtcagcgtgggagcgctacgtggccatccta 438
            |||     ||  |  |||||||||  | || || |        | || |||||||||||||
Db  458 TGCACCGCCTCCATCTTCAACCTGTGCGCCATCAGCGTGGACAGGTTCGTGGCCGTGGCC 517

Qy  439 cacccgttccgcgccaaactgcaagagacaccggcgggccctcaggatcctcggcgatc 498
            || | ||||| |||| || || || |  ||| ||  |  ||||||||| |   |||||||
Db  518 GTGCCGCTGCCCTACAACCGGCAGAGGTGGGAGCCGCAGCTGCTCATCATCGTGTCGTGGGCC 577

Qy  499 gtctgggggcttctccgtgtctcttccctgcccaacaccagcatccaatggcatcaaagttc 558
            | || |   |  |  |  |  |  |||   |||| ||| ||||| |  |||||||| |||
Db  578 ACGTGGCTGCTGTCCGGCCGGCGGCGGGGTTGGGGCGGCCCCTACTCGGCCTCAACGACGTGGCC 637

Qy  559 cactacttccccaatgggtccc 580
            |   |   ||| |||||||||
Db  638 GGCCCGGACCCGCCCGTGTGCC 659
```

RESULT 14
US-07-928-611-19
Sequence 19, Application US/07928611
Patent No. 5569601
GENERAL INFORMATION:
  APPLICANT: Van Tol, Hubert H.M.
  APPLICANT: Civelli, Olivier
  TITLE OF INVENTION: A NO. 5569601el Human Dopamine Receptor and Uses
  NUMBER OF SEQUENCES: 22
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Allegretti & Witcoff, Ltd.
    STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/07/928,611
  FILING DATE: 19920810
  CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
  NAME: No. 5569601nan, Kevin E
  REGISTRATION NUMBER: 35,303
  REFERENCE/DOCKET NUMBER: 90,1092-B
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 312-715-1000
  TELEFAX: 312-715-1234
  TELEX: 810-221-8317
INFORMATION FOR SEQ ID NO: 19:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1466 base pairs
    TYPE: NUCLEIC ACID
    STRANDEDNESS: single
    TOPOLOGY: linear
  MOLECULE TYPE: cDNA
  FEATURE:
    NAME/KEY: 5'UTR
    LOCATION: 1..103
  FEATURE:
    NAME/KEY: 3'UTR
    LOCATION: 1364..1466
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 104..1363
US-07-928-611-19

Query Match   15.4%;   Score 112.4;   DB 1;   Length 1466;
Best Local Similarity   53.4%;   Pred. No. 8.5e-16;
Matches 236;   Conservative 0;   Mismatches 206;   Indels 0;   Gaps 0;

```
Qy  139 gtggtgtatgtgccaatttttgtgtggggtcattggcaatgtcctggttgtgcctggtg 198
            ||||                  |  |   |||                ||| |||||
Db  218 GTGGGGGGGCGTGCTGTCTCATCGGCCCGGTGCTCGCGGGGGAACTCGCTCGTGCGGTGAGC 277

Qy  199 atctgcagcacagggctatgaagagacgcccacacaactactactcttcagcctggcgtc 258
            |  |   |   | |   |  ||  |   |  | |||||   ||||||||||||||||
Db  278 GTGGCCACCGAGCGCGCCCCTGCAGACGCCCACCAACTCCTTCATCGTGACGCTGGGCGGC 337

Qy  259 tctgacctcctggtctctgctcctgggaatgccctgtctatgagagtgtggcgcaac 318
            |  |||||||| |  |      |  |||||||  |||||||||||||  ||||||
Db  338 GCCGACCTCCTCCTCCTGGTGCTCGCGCGTCTTCGTCTACTCCGAGGTCCAGGT 397

Qy  319 taccctttcttgttcgggccgtgggctgctacttcaagacggccctctctttgagaccgtg 378
            |||   |    |  ||| |||  |  |||  |        | ||   | ||||||||||||
Db  398 GGGCGGCGTGGCTGCTGAGCCCCGGCGCCCTGCCCGCCCATGGCCATGGACGTCATGCTG 457

Qy  379 tgcttccgcctccatccagcatcaccacacgtcagcgtgggagcgctacgtggccatccta 438
            |||     ||  |  |||||||||  | || || |        | || |||||||||||||
Db  458 TGCACCGCCTCCATCTTCAACCTGTGCGCCATCAGCGTGGACAGGTTCGTGGCCGTGGCC 517

Qy  439 cacccgttccgcgccaaactgcaagagacaccggcgggccctcaggatcctcggcgatc 498
            || | ||||| |||| || || || |  ||| ||  |  ||||||||| |   |||||||
Db  518 GTGCCGCTGCCCTACAACCGGCAGAGGTGGGAGCCGCAGCTGCTCATCATCGTGTCGTGGGCC 577

Qy  499 gtctgggggcttctccgtgtctcttccctgcccaacaccagcatccaatggcatcaaagttc 558
            | || |   |  |  |  |  |  |||   |||| ||| ||||| |  |||||||| |||
Db  578 ACGTGGCTGCTGTCCGGCCGGCGGCGGGGTTGGGGCGGCCCCTACTCGGCCTCAACGACGTGGCC 637

Qy  559 cactacttccccaatgggtccc 580
```

Db   638 GGCCGCGGACCCCGCCGCCGTGTGCC 659

RESULT 15
US-08-487-811A-19
Sequence 19, Application US/08487811A
Patent No. 5883226
GENERAL INFORMATION:
APPLICANT: Civelli, Olivier
APPLICANT: Van Tol, Hubert H.M.
TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,811A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: No. 5883226nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092-L
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX:
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1466 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..103
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 1364..1466
FEATURE:
NAME/KEY: CDS
LOCATION: 104..1363
US-08-487-811A-19

Query Match   15.4%;   Score 112.4;   DB 2;   Length 1466;
Best Local Similarity 53.4%;   Pred. No. 8.5e-16;   DB 206;
Matches 236;   Conservative   0;   Mismatches   206;   Indels   0;   Gaps   0;

Qy   139 gtggtgtatgtgccaattttggtggggtcattggcaatgtcctgtgtgcctggtg 198
         ||||| |||| ||||| ||| |||||||| ||| || ||||||| ||| |||||
Db   218 GTGGGGGGGCGTGCTGCTCATCGGCGGGAACTCGCTCGTTCGTGTGCGTGAGC 277

Qy   199 atttctgcagcaccaggctatgaagacgcccacaacaactactacctcttcagcctggcggtc 258
         ||| | |||| ||||| | || ||| ||| || || | || |||||||| ||||| ||||| |
Db   278 GTGGCCACGAGCGGCGCCCTGCAGACGCCCACAACTCCTTCATCGTCGACTCGGCCTGGCGGCC 337

Qy   259 tctgacctcctggtcctgctccttggaatgcccctggaggtctatgagatggccgcaac 318
         | |||||||||| | |||| ||| || || |||| |||||| ||| ||| || |||
Db   338 GCCGACCTCCTCCTCGCTCTCCTTCGTTCGGTCGCTCTCGGCGACGTGCCAGGGT 397

Qy   319 tacccttcttgttcgggcccgtgggctgctacttcaagacacggccctcttgagacacgtg 378
         | ||||| ||||||| | ||| ||| |||| | ||| ||||| || ||| |||| |
Db   398 GGCGGCGTTGGGCTTGCTGCTGGAGCCCCCGCCGCCTTGTGGACGCGCCCTTCATGGCCATGGAGCGTCATCGCTG 457

Qy   379 tgcttcgccctccaatcctcagcatcaccaccgtcagcgtgggagcgctacgtggccatccta 438
         ||| | ||||||| || |||||| ||||||||| ||| |||| || || ||||||| ||
Db   458 TGCACCGCGCCTCCATCTTCAAACCTGTGCCCATCAGCGTGGACAGGGTTCGTGCCGTGGCC 517

Qy   439 cacccgttccgcgccaaactgcaagagaccaccggacgccggggaccctcagggatcctcggcatc 498
         || |||||| | || || |||| ||| || |||| || ||||| || ||| ||| |||| ||||
Db   518 GTGCCGCGTGCCGTACAACCGGCACGAGCTGGGAGCCGCCCGGCAGCTGCTGCTCATCGCGCC 577

Qy   499 gtctggggctctccgtgctctcccctgcccaacaccagcatccaaggcatcaagttc 558
         || ||| || ||||| |||||| ||| ||| || || ||||| ||| |||| |
Db   578 ACGTGCGCTGCTCTGTCTCCGGGGGCCGTGGCGCGGGCGGGCCGCGGGGCCGGG 637

Qy   559 cactacttcccccaatggggtccc 580
         || |||||||| || ||| ||
Db   638 GGCCGCGGACCCCGCCGCCGTGTGCC 659

Search completed: August 30, 2001, 02:47:22
Job time: 3552 sec